Chapter 4 Routing Protocol RelatedConfiguration 3

Chapter 4 Routing Protocol Related Configuration

4.1 Routing Protocol Overview

To communicate with a remote host over the Internet, a host must choose a proper route via a set of routers or Layer3 switches.

Both routers and layer3 switches calculate the route using CPU, the difference is that layer3 switch adds the calculated route to the switch chip and forward by the chip at wire speed, while the router always store the calculated route in the route table or route buffer, and data forwarding is performed by the CPU. For this reason, although both routers and switches can perform route selection, layer3 switches have great advantage over routers in data forwarding. The following describes basic principle and methods used in layer3 switch route selection.

In route selection, the responsibility of each layer3 switch is to select a proper midway route according to the destination of the packet received; and send the packet to the next layer3 switch until the last layer3 switch in the route send the packet to the destination host. A route is the path selected by each layer3 switch to pass the packet to the next layer3 switch. Route can be grouped into direct route, static route and dynamic route.

Direct route refer to the path directly connects to the layer3 switch, and can be obtained with no calculation.

Static route is the manually specified path to a network or a host; static route cannot be changed freely. The advantage of static route is simple and consistent, and it can limit illegal route modification, and is convenient for load balance and route backup. However, as this is set manually, it is not suitable for mid- or large-scale networks for the route in such conditions are too huge and complex.

Dynamic route is the path to a network or a host calculated by the layer3 switch according to the routing protocols enabled. If the next hop layer3 switch in the path is not reachable, layer3 switch will automatically discard the path to that next hop layer3 switch and choose the path through other layer3 switches.

There are two dynamic routing protocols: Interior Gateway Protocol (IGP) and Exterior Gateway protocol (EGP). IGP is the protocol used to calculate the route to a destination inside an autonomous system. IGP supported by switch include RIP and OSPF, RIP and OSRF can be configured according to the requirement. Switch supports running several IGP dynamic routing protocols at the same time. Or, other dynamic routing protocols and static route can be introduced to a dynamic routing protocol, so that multiple routing protocols can be associated.

EGP is used to exchange routing information among different autonomous systems, such as BGP protocol. EGP supported by switch include BGP-4, BGP-4+.

4.1.1 Routing Table

As mentioned before, layer3 switch is mainly used to establish the route from the current layer3 switch to a network or a host, and to forward packets according to the route. Each layer3 switch has its own route table containing all routes used by that switch. Each route entry in the route table specifies the physical port should be used for forwarding packet to reach a destination host or the next hop layer3 switch to the host.

The route table mainly consists of the following:

- ☞ Destination address: used to identify the destination address or destination network of an IP packet.
- ☞ Network mask: used together with destination address to identify the destination host or the network the layer3 switch resides. Network mask consists of several consecutive binary 1's, and usually in the format of dotted decimal (an address consists of 1 to 4 255's.) When "AND" the destination address with network mask, we can get the network address for the destination host or the network the layer3 switch resides. For example, the network address of a host or the segment the layer3 switch resides with a destination address of 200.1.1.1 and mask 255.255.255.0 is 200.1.1.0.
- ☞ Output interface: specify the interface of layer3 switch to forward IP packets.
- ☞ IP address of the next layer3 switch (next hop): specify the next layer3 switch the IP packet will pass.
- ☞ Route entry priority: There may be several different next hop routes leading to the same destination. Those routes may be discovered by different dynamic routing protocols or static routes manually configured. The entry with the highest priority (smallest value) becomes the current best route. The user can configure several routes of different priority to the same destination; layer3 switch will choose one route for IP packet forwarding according to the priority order.

To prevent too large route table, a default route can be set. Once route table look up fails, the default route will be chosen for forwarding packets.

The table below describes the routing protocols supported by switch and the default route look up priority value.

Routing Protocols or route type	Default priority value
Direct route	0
OSPF	110
Static route	1
RIP	120
OSPF ASE	150
IBGP	200
EBGP	20
Unknown route	255

4.1.2 IP Routing Policy

4.1.2.1 Introduction to Routing Policy

Some policies have to be applied when the router publishing and receiving routing messages so to filter routing messages, such as only receiving or publishing routing messages meets the specified conditions. A routing protocol maybe need redistribute other routing messages found by other protocols such as OSPF so to increase its own routing knowledge; when the router redistributing routing messages from other routing protocols there may be only part of the qualified routing messages is needed, and some properties may have to be configured to suit this protocol.

To achieve routing policy, first we have to define the characteristics of the routing messages to be applied with routing policies, namely define a group matching rules. We can configure by different properties in the routing messages such as destination address, the router address publishing the routing messages. The matching rules can be previously configured to be applied in the routing publishing, receiving and distributing policies.

Five filters are provided in switch: route-map, acl, as-path, community-list and ip-prefix for use. We will introduce each filter in following sections:

1. route-map

For matching certain properties of the specified routing information and setting some routing propertities when the conditions are fulfilled.

Route-map is for controlling and changing the routing messages while also controlling the redistribution among routes. A route-map consists of a series of match and set commands in which the match command specifies the conditions required matching, and the set command specifies the actions to be taken when matches. The route-map is also for controlling route publishing among different route process. It can also used on policy routing which select different routes for the messages other than the shortest route.

A group matches and set clauses make up a node. A route-map may consist of several nodes each of which is a unit for matching test. We match among nodes with by sequence-number. Match clauses define matching rules. The matching objects are some properties of routing messages. Different match clause in the same node is “and” relation logically, which means the matching test of a node, will not be passed until conditions in its entire match clause are matched. Set clause specifies actions, namely configure some properties of routing messages after the matching test is passed.

Different nodes in a route-map is an “or” relation logically. The system checks each node of the route-map in turn and once certain node test is passed the route-map test will be passed without taking the next node test.

2. access control list(acl)

ACL (Access Control Lists) is a data packet filter mechanism in the switch. The switch controls the network access and secure the network service by permitting or denying certain data packet transmtting out from or into the network. Users can establish a group of rules by certain messages in the packet, in which each rule to be applied on certain amount of matching messages: permit or deny. The users can apply these rules to the entrance or exit of specified switch, with which data stream in certain direction on certain port would have to follow the specified ACL rules in-and-out the switch. Please refer to chapter “ACL Configuration”.

3. Ip-prefix list

The ip-prefix list acts similarly to acl while more flexible and more understandable. The match object of ip-prefix is the destination address messages field of routing messages when applied in routing messages filtering.

An ip-prefix is identified by prefix list name. Each prefix list may contain multiple items, each of which specifies a matching range of a network prefix type and identifies with a sequence-number which specifies the matching check order of ip-prefix.

In the process of matching, the switch check each items identified by sequence-number in ascending order and the filter will be passed once certain items is matched(without checking rest items)

4. Autonomic system path information access-list as-path

The autonomic system path information access-list as-path is only used in BGP. In the BGP routing messages packet there is an autonomic system path field (in which autonomic system path the routing messages passes through is recorded). As-path is specially for specifying matching conditions for autonomic system path field.

As for relevant as-path configurations, please refer to the ip as-path command in BGP configuration.

5. community-list

Community-list is only for BGP. There is a community property field in the BGP routing messages packet for identifying a community. The community list is for specifying matching conditions for Community-list field.

As for relevant Community-list configuration, please refer to the ip as-path command in BGP configuration

4.1.2.2 IP Routing Policy Configuration Task List

1. Define route-map
2. Define the match clause in route-map
3. Define the set clause in route-map
4. Define address prefix list

1. Define route-map

Command	Explanation
Global mode	
route-map <*map_name*> {deny | permit} <*sequence_num*> **no route-map <*map_name*> [{deny | permit} <*sequence_num*>]**	Configure route-map; the **no route-map <*map_name*> [{deny | permit} <*sequence_num*>]** command deletes the route-map.

2. Define the match clause in route-map

Command	Explanation
Route-map configuration mode	

match as-path <*list-name*> **no match as-path [<*list-name*>]**	Match the autonomous system as path access-list the BGP route passes through; the **no match as-path [<*list-name*>]** command deletes match condition.
**match community <*community-list-name*	*community-list-num* > [exact-match]** **no match community [<*community-list-name*	*community-list-num* > [exact-match]]**	Match a community property access-list. The **no match community [<*community-list-name*	*community-list-num* > [exact-match]]** command deletes match condition.
match interface <*interface-name* > **no match interface [<*interface-name* >]**	Match by ports; The **no match interface [<*interface-name* >]** command deletesmatch condition.
**match ip <address	next-hop> <*ip-acl-name*	*ip-acl-num*	prefix-list *list-name*>** **no match ip <address	next-hop> [<*ip-acl-name*	*ip-acl-num*	prefix-list [*list-name*]>]**	Match the address or next-hop; The **no match ip <address	next-hop> [<*ip-acl-name*	*ip-acl-num*	prefix-list [*list-name*]>]** command deletes match condition.
match metric <*metric-val* > **no match metric [<*metric-val* >]**	Match the routing metric value; The **no match metric [<*metric-val* >]** command deletes match condition.
**match origin <egp	igp	incomplete >** **no match origin [<egp	igp	incomplete >]**	Match the route origin; The **no match origin [<egp	igp	incomplete >]** command deletes match condition.
**match route-type external <type-1	type-2 >** **no match route-type external [<type-1	type-2 >]**	Match the route type; The **no match route-type external [<type-1	type-2 >]** command deletes match condition.
match tag <*tag-val* > **no match tag [<*tag-val* >]**	Match the route tag; The **no match tag [<*tag-val* >]** command deletes match condition.

3. Define the set clause in route-map

Command	Explanation
Route-map configuration mode	
set aggregator as <*as-number*> <*ip_addr*> **no set aggregator as [<*as-number*> <*ip_addr*>]**	Distribute an AS No. for BGP aggregator; The no command deletes the configuration
set as-path prepend <*as-num*> **no set as-path prepend [<*as-num*>]**	Add a specified AS No. before the BGP routing messages as-path series; The no command deletes the configuration
set atomic-aggregate **no set atomic-aggregate**	Configure the BGP atomic aggregate property; The no command deletes the configuration
set comm-list <*community-list-name* | *community-list-num* > delete **no set comm-list <*community-list-name* | *community-list-num* > delete**	Delete BGP community list value; The no command deletes the configuration
set community [*AA:NN*] [internet] [local-AS] [no-advertise] [no-export] [none] [additive] **no set community [*AA:NN*] [internet] [local-AS] [no-advertise] [no-export] [none] [additive]**	Configure BGP community list value; The no command deletes the configuration
set extcommunity <rt | soo> <*AA:NN*> **no set extcommunity <rt | soo> [<*AA:NN*>]**	Configure BGP extended community list property; The no command deletes the configuration
set ip next-hop <*ip_addr*> **no set ip next-hop [<*ip_addr*>]**	Set next-hop IP address; The no command deletes the configuration
set local-preference <*pre_val*> **no set local-preference [<*pre_val*>]**	Set local preference; The no command deletes the configuration
set metric < *+/- metric_val | metric_val*> **no set metric [*+/- metric_val | metric_val*]**	Set routing metric value; The no command deletes the configuration
set metric-type <type-1 | type-2> **no set metric-type [<type-1 | type-2>]**	Set OSPF metric type; The no command deletes the configuration
set origin <egp | igp | incomplete > **no set origin [<egp | igp | incomplete >]**	Set BGP routing origin; The no command deletes the configuration

set originator-id <*ip_addr*> **no set originator-id [<*ip_addr*>]**	Set routing originator ID; The no command deletes the configuration
set tag <*tag_val*> **no set tag [<*tag_val*>]**	Set OSPF routing tag value; The no command deletes the configuration
set vpnv4 next-hop <*ip_addr*> **no set vpnv4 next-hop [<*ip_addr*>]**	Set BGP VPNv4 next-hop address; the no command deletes the configuration
set weight < *weight_val*> **no set weight [<*weight_val*>]**	Set BGP routing weight; The no command deletes the configuration

4. Define address prefix list

Command	Explanation
Global mode	
ip prefix-list <*list_name*> description <*description*> **no ip prefix-list <*list_name*> description**	Describe the prefix list; The **no ip prefix-list <*list_name*> description** command deletes the configuration.
ip prefix-list <*list_name*> [seq <*sequence_number*>] <deny | permit> < any / *ip_addr/mask_length* [ge *min_prefix_len*] [le *max_prefix_len*]> **no ip prefix-list <*list_name*> [seq <*sequence_number*>] [<deny | permit> < any / *ip_addr/mask_length* [ge *min_prefix_len*] [le *max_prefix_len*]>]**	Set the prefix list; The **no ip prefix-list <*list_name*> [seq <*sequence_number*>] [<deny | permit> < any / *ip_addr/mask_length* [ge *min_prefix_len*] [le *max_prefix_len*]>]** command deletes the configuration.

4.1.2.3 Configuration Examples

The figure below shows a network consisting of four Layer 3 switches. This example demonstrates how to set the BGP as-path properties through route-map. BGP protocol is applied among the Layer 3 switches. As for SwitchC, the network 192.68.11.0/24 can be reached through two paths in which one is AS-PATH 1 by IBGP (going through SwitchD), the other one is AS-PATH 2 by EBGP (going through SwitchB). BGP selects the shortest path, so AS-PATH 1 is the preferred path. If the path 2 is wished, which is through EBGP path, we can add two extra AS path numbers into the AS-PATH messages from SwitchA to SwitchD so as to change the determination SwitchC take to 192.68.11.0/24.

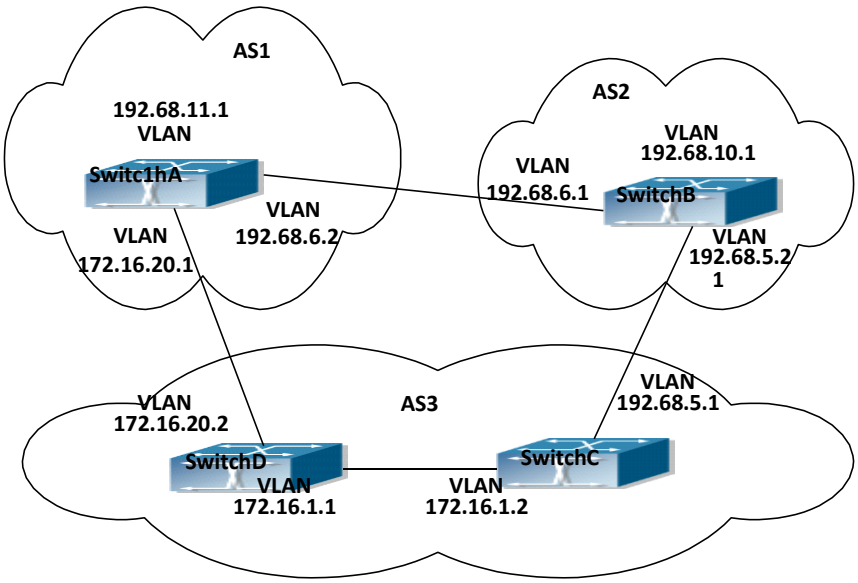

Figure 4-1 Policy routing Configuration

Configuration procedure: (only SwitchA is listed, configurations for other switches are omitted.)

The configuration of Layer 3 switchA:

SwitchA#config

SwitchA(config) #router bgp 1

SwitchA(config-router)#network 192.68.11.0 mask 255.255.255.0

SwitchA(config-router)#neighbor 172.16.20.2 remote-as 3

SwitchA(config-router)#neighbor 172.16.20.2 route-map AddAsNumbers out

SwitchA(config-router)#neighbor 192.68.6.1 remote-as 2

SwitchA(config-router)#exit

SwitchA(config)#route-map AddAsNumbers permit 10

SwitchA(config-route-map)#set as-path prepend 1 1

4.1.2.4 Troubleshooting

Faq: The routing protocol could not achieve the routing messages study under normal protocol running state

Troubleshooting: check following errors:

☞ Each node of route-map should at least has one node is permit match mode. When the route map is used in routing messages filtering, the routing messages will be considered not pass the routing messages filtering if certain routing messages does not pass the filtering of any nodes. When all nodes are set to deny mode, all routing messages will not pass the filtering in this route-map.

☞ Items in address prefix list should at least have one item set to permit mode. The deny mode items can be defined first to fast remove the unmatched routing messages, however if all the items are set to deny mode, any route will not be able to pass the filtering of this

address prefix list. We can define a permit 0.0.0.0/0 le 32 item after several deny mode items are defined so to permit all other routing messages pass through. Only default route will be matched in less-equal 32 is not specified.

4.2 Static Route

4.2.1 Introduction to Static Route

As mentioned earlier, the static route is the manually specified path to a network or a host. Static route is simple and consistent, and can prevent illegal route modification, and is convenient for load balance and route backup. However, it also has its own defects. Static route, as its name indicates, is static, it won't modify the route automatically on network failure, and manual configuration is required on such occasions, therefore it is not suitable for mid and large-scale networks.

Static route is mainly used in the following two conditions: 1) in stable networks to reduce load of route selection and routing data streams. For example, static route can be used in route to STUB network. 2) For route backup, configure static route in the backup line, with a lower priority than the main line.

Static route and dynamic route can coexist; layer3 switch will choose the route with the highest priority according to the priority of routing protocols. At the same time, static route can be introduced (redistribute) in dynamic route, and change the priority of the static route introduced as required.

4.2.2 Introduction to Default Route

Default route is a kind of static route, which is used only when no matching route is found. In the route table, default route in is indicated by a destination address of 0.0.0.0 and a network mask of 0.0.0.0, too. If the route table does not have the destination of a packet and has no default route configured, the packet will be discarded, and an ICMP packet will be sent to the source address indicate the destination address or network is unreachable.

4.2.3 Static Route Configuration Task List

1. Static route configuration
2. VRF configuration

1. Static route configuration

Command	Explanation
Global mode	

ip route {*<ip-prefix>* *<mask>* | *<ip-prefix>/<prefix-length>*} {*<gateway-address>* | *<gateway-interface>*} [*<distance>*] **no ip route {*<ip-prefix>* *<mask>* | *<ip-prefix>/<prefix-length>*} [*<gateway-address>* | *<gateway-interface>*] [*<distance>*]**	Set static routing; the **no ip route {*<ip-prefix>* *<mask>* | *<ip-prefix>/<prefix-length>*} [*<gateway-address>* | *<gateway-interface>*] [*<distance>*]** command deletes a static route entry

2. VRF configuration

Command	Explanation
Global mode	
ip route vrf <name> {<ip-prefix> <mask>|<ip-prefix/<prefix-length>} {<gateway-address>|<gateway-interface>} [<distance>] **no ip route vrf <name> {<ip-prefix> <mask>|<ip-prefix/<prefix-length>} [<gateway-address>|<gateway-interface>] [<distance>]**	Configure the static route, the no command will delete the static route.

4.2.4 Static Route Configuration Examples

The figure shown below is a simple network consisting of three layer3 switches, the network mask for all switches and PC is 255.255.255.0. PC-A and PC-C are connected via the static route set in SwtichA and SwitchC; PC3 and PC-B are connected via the static route set in SwtichC to SwitchB; PC-B and PC-C is connected via the default route set in SwitchB.

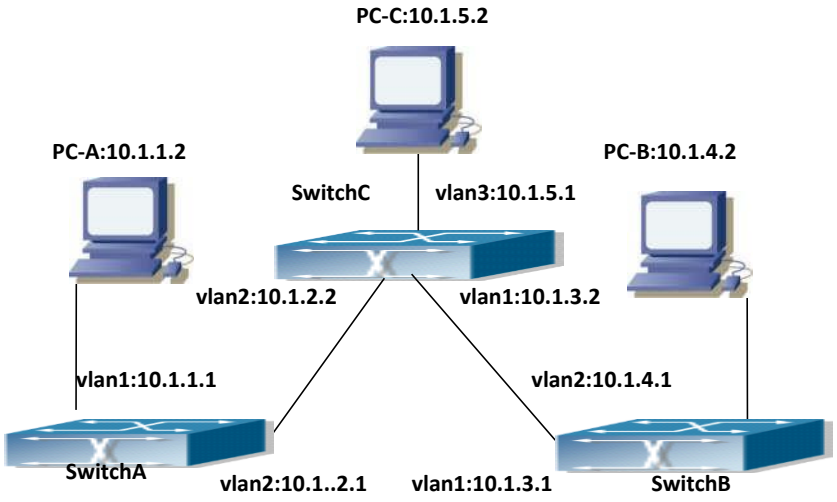

Figure 4-2 Static Route Configurations

Configuration steps:

Configuration of layer3 SwitchA

Switch#config

Switch (config) #ip route 10.1.5.0 255.255.255.0 10.1.2.2
Configuration of layer3 SwitchC
Switch#config
Next hop use the partner IP address
Switch(config)#ip route 10.1.1.0 255.255.255.0 10.1.2.1
Next hop use the partner IP address
Switch(config)#ip route 10.1.4.0 255.255.255.0 10.1.3.1
Configuration of layer3 SwitchB
Switch#config
Switch(config)#ip route 0.0.0.0 0.0.0.0 10.1.3.2
In this way, ping connectivity can be established between PC-A and PC-C, and PC-B and PC-C.

4.3 RIP

4.3.1 Introduction to RIP

RIP is first introduced in ARPANET, this is a protocol dedicated to small, simple networks. RIP is a distance vector routing protocol based on the Bellman-Ford algorithm. Network devices running vector routing protocol send two kind of information to the neighboring devices regularly:

• Number of hops to reach the destination network, or metrics to use or number of networks to pass.

• What is the next hop, or the director (vector) to use to reach the destination network.

The distance vector Layer 3 switch send all their route selecting tables to the neighbor layer3 switches at regular interval. A layer3 switch will build their own route selecting information table based on the information received from the neighbor layer3 switches. Then, it will send this information to its own neighbor layer3 switches. As a result, the route selection table is built on second hand information, route beyond 15 hops will be deemed as unreachable.

RIP protocol is an optional routing protocol based on UDP. Hosts using RIP send and receive packets on UDP port 520. All layer3 switches running RIP send their route table to all neighbor layer3 switches every 30 seconds for update. If no information from the partner is received in 180 seconds, then the device is deemed to have failed and the network connected to that device is considered to be unreachable. However, the route of that layer3 switch will be kept in the route table for another 120 seconds before deletion.

As layer3 switches running RIP built route table with second hand information, infinite count may occur. For a network running RIP routing protocol, when an RIP route becomes unreachable, the neighboring RIP layer3 switch will not send route update packets at once, instead, it waits until the update interval timeout (every 30 seconds) and sends the update packets containing that route. If before it receives the updated packet, its neighbors send packets containing the information about the failed neighbor, “infinite count” will be resulted. In other words, the route

of unreachable layer3 switch will be selected with the metrics increasing progressively. This greatly affects the route selection and route aggregation time.

To prevent “infinite count”, RIP provides mechanism such as “split horizon” and “triggered update” to solve route loop. “Split horizon” is done by avoiding sending to a gateway routes leaned from that gateway. There are two split horizon methods: “simple split horizon” and “poison reverse split horizon”. Simple split horizon deletes from the route to be sent to the neighbor gateways the routes learnt from the neighbor gateways; poison reverse split horizon not only deletes the abovementioned routes, but set the costs of those routes to infinite. “Triggering update” mechanism defines whenever route metric changed by the gateway, the gateway advertise the update packets immediately, regardless of the 30 second update timer status.

There two versions of RIP, version 1 and version 2. RFC1058 introduces RIP-I protocol, RFC2453 introduces RIP-II, which is compatible with RFC1723 and RFC1388. RIP-I updates packets by packets broadcast, subnet mask and authentication is not supported. Some fields in the RIP-I packets are not used and are required to be all 0’s; for this reason, such all 0's fields should be checked when using RIP-I, the RIP-I packets should be discarded if such fields are non-zero. RIP-II is a more improved version than RIP-I. RIP-II sends route update packets by multicast packets (multicast address is 224.0.0.9). Subnet mask field and RIP authentication filed (simple plaintext password and MD5 password authentication are supported), and support variable length subnet mask. RIP-II used some of the zero field of RIP-I and require no zero field verification. switch send RIP-II packets in multicast by default, both RIP-I and RIP-II packets will be accepted.

Each layer3 switch running RIP has a route database, which contains all route entries for reachable destination, and route table is built based on this database. When a RIP layer3 switch sent route update packets to its neighbor devices, the complete route table is included in the packets. Therefore, in a large network, routing data to be transferred and processed for each layer3 switch is quite large, causing degraded network performance.

Besides the above mentioned, RIP protocol allows route information discovered by the other routing protocols to be introduced to the route table. It can also be as the protocol exchanging route messages with CE on PE routers, and supports the VPN route/transmitting examples.

The operation of RIP protocol is shown below:

1. Enable RIP. The switch sends request packets to the neighbor layer3 switches by broadcasting; on receiving the request, the neighbor devices reply with the packets containing their local routing information.
2. The Layer3 switch modifies its local route table on receiving the reply packets and sends triggered update packets to the neighbor devices to advertise route update information. On receiving the triggered update packet, the neighbor lay3 switches send triggered update packets to their neighbor lay3 switches. After a sequence of triggered update packet broadcast, all layer3 switches get and maintain the latest route information.

In addition, RIP layer3 switches will advertise its local route table to their neighbor devices every 30 seconds. On receiving the packets, neighbor devices maintain their local route table, select the best route and advertise the updated information to their own neighbor devices, so that the updated routes are globally valid. Moreover, RIP uses a timeout mechanism for outdated route, that is, if a switch does not receive regular update packets from a neighbor within a certain

interval (invalid timer interval), it considers the route from that neighbor invalid, after holding the route fro a certain interval (holddown timer interval), it will delete that route.

4.3.2 RIP Configuration Task List

1. Enable RIP (required)
   (1) Enable/disable RIP module.
   (2) Enable interface to send/receive RIP packets
2. Configure RIP protocol parameters (optional)
   (1) Configure RIP sending mechanism
      1) Configure specified RIP packets transmission address
      2) Configure RIP interface broadcast
   (2) Configure the RIP routing parameters
      1) Configure route introduction (default route metric, configure routes of the other protocols to be introduced in RIP)
      2) Configure interface authentication mode and password
      3) Configure the route deviation
      4) Configure and apply route filter
      5) Configure Split Horizon
   (3) Configure other RIP protocol parameters
      1) Configure the managing distance of RIP route
      2) Configure the RIP route capacity limit in route table
      3) Configure the RIP update, timeout, holddown and other timer.
      4) Configure the receiving buffer size of RIP UDP
3. Configure RIP-I/RIP-II switch
   (1) Configure the RIP version to be used in all interfaces
   (2) Configure the RIP version to send/receive in all interfaces
   (3) Configure whether to enable RIP packets sending/receiving for interfaces
4. Delete the specified route in RIP route table
5. Configure the RIP routing aggregation
   (1) Configure aggregation route of IPv4 route mode
   (2) Configure aggregation route of IPv4 interface configuration mode
   (3) Display IPv4 aggregation route information
6. Configure redistribution of OSPF routing to RIP
   (1) Enable Redistribution of OSPF routing to RIP
   (2) Display and debug the information about configuration of redistribution of OSPF routing to RIP
7. Configure VRF address family mode for RIP
   (1) Enable/disable RIP module
   (2) Configure VRF address family

1. Enable RIP protocol

Applying RIP route protocol with basic configuration in switch is simple. Normally you only

have to open the RIP switch and configure the segments running RIP, namely send and receive the RIP data packet by default RIP configuration. The version of data packet sending and receiving is variable when needed, allow/deny sending, receiving RIP data packet. Refer to 3.

Command	Explanation
Global Mode	
router rip **no router rip**	Enables RIP; the **no router rip** command disables RIP.
Router and address family configuration mode	
**network <*A.B.C.D/M*	*ifname|vlan*>** **no network <*A.B.C.D/M*	*ifname|vlan*>**	Enables the segment running RIP protocol; the **no network <*A.B.C.D/M*	*ifname|vlan*>** command deletes the segment.

2. Configure RIP protocol parameters

（1）Configure RIP packet transmitting mechanism

1) Configure the RIP data packet point-transmitting

2) Configure the Rip broadcast

Command	Explanation
Router Configuration Mode	
neighbor <*A.B.C.D*> **no neighbor <*A.B.C.D*>**	Specify the IP address of the neighbor router needs point-transmitting; the **no neighbor <*A.B.C.D*>** command cancels the appointed router.
passive-interface<*ifname|vlan*> **no passive-interface<*ifname|vlan* >**	Block the RIP broadcast on specified pot and the RIP data packet is only transmittable among Layer 3 switch configured with neighbor. The **no passive-interface<*ifname|vlan* >** command cancels the function.

（2）Configure RIP route parameters

1） Configure route introduction (default route metric, configure routes of the other protocols to be introduced in RIP)

Command	Explanation
Router Configuration Mode	
default-metric <*value*> **no default-metric**	Sets the default route metric for route to be introduced; the **no default-metric**command restores the default setting.
**redistribute {kernel |connected| static| ospf	isis| bgp} [metric<*value*>] [route-map<*word*>]** **no redistribute {kernel |connected| static| ospf	isis| bgp} [metric<*value*>] [route-map<*word*>]**	Redistribute the routes distributed in other routing protocols into the RIP data packet; the **no redistribute {kernel |connected| static| ospf	isis| bgp} [metric<value>] [route-map<word>]** command cancels the distributed route of corresponding protocols.

default-information originate **no default-information originate**	Generate a default route to the RIP protocol; the **no default-information originate** command cancels the feature.

2） Configure interface authentication mode and password

Command	Explanation
Interface configuration mode	
ip rip authentication mode { text| md5} **no ip rip authentication mode [text| md5]**	Sets the authentication method; the **no ip rip authentication mode [text| md5]** command cancels the authentication action.
ip rip authentication string *<text>* **no ip rip authentication string**	Sets the authentication key; the **no ip rip authentication string** command means no key is needed.
ip rip authentication key-chain *<name-of-chain>* **no ip rip authentication key-chain [*<name-of-chain>*]**	Sets the key chain used in authentication, the **no ip rip authentication key-chain [*<name-of-chain>*]** command means the key chain is not used.
ip rip authentication cisco-compatible **no ip rip authentication cisco-compatible**	After configure this command, configure MD5 authentication, then can receive RIP packet of Cisco, the no command restores the default configuration.
Global mode	
key chain *<name-of-chain>* **no key chain *< name-of-chain >***	Enter keychain mode, and configure a key chain, the **no key chain *< name-of-chain >*** command deletes the key chain.
Keychain mode	
key *<keyid>* **no key *<keyid>***	Enter the keychain-key mode and configure a key of the keychain; the **no key *<keyid>*** command deletes one key.
Keychain-key mode	
key-string *<text>* **no key-string *<text>***	Configure the password used by the key, the **no key-string *<text>*** command deletes the password.
accept-lifetime *<start-time>* {*<end-time>*| **duration*<seconds>*| infinite}** **no accept-lifetime**	Configure a key on the key chain and accept it as an authorized time; the **no accept-lifetime** command deletes it.
send-lifetime *<start-time>* {*<end-time>*| **duration*<seconds>*| infinite}** **no send-lifetime**	Configure the transmitting period of a key on the key chain; the **no send-lifetime** command deletes the send-lifetime.

3） Configure the route deviation

Command	Explanation
Router configuration mode	

offset-list <access-list-number | access-list-name> {in | out } <number> [<ifname>] **no offset-list <access-list-number |access-list-name> {in|out }<number >[<ifname>]**	Configure that provide a deviation value to the route metric value when the port sends or receives RIP data packet; the **no offset-list <access-list-number |access-list-name> {in|out } <number >[<ifname>]** command removes the deviation table.

4） Configure and apply the route filtering

Command	Explanation
Router configuration mode	
distribute-list {< *access-list-number /access-list-name* >|prefix<*prefix-list-name*>}{in|out} [<*ifname*>] **no distribute-list {< *access-list-number /access-list-name* >|prefix<*prefix-list-name*>}{in|out} [<*ifname*>]**	Configure and apply the access table and prefix table to filter the routes. The **no distribute-list {< *access-list-number /access-list-name*>|prefix<*prefix-list-name*>}{in|out} [<*ifname*>]** command means do not use the access table and prefix table.

5） Configure the split horizon

Command	Explanation
Interface configuration mode	
ip rip split-horizon [poisoned] **no ip rip split-horizon**	Configure that take the split horizon when the port sends data packets; poisoned for poison reverse the **no ip rip split-horizon** command cancels the split horizon.

（3） Configure other RIP protocol parameters

1） Configure RIP routing priority

2） Configure the RIP route capacity limit in route table

3） Configure timer for RIP update, timeout and hold-own

4） Configure RIP UDP receiving buffer size

Command	Explanation
Router configuration mode	
distance <*number*> [<*A.B.C.D/M*>] [<*access-list-name|access-list-number* >] **no distance [<*A.B.C.D/M*>]**	Specify the route administratively distance of RIP protocol; the **no distance [<*A.B.C.D/M*>]** command restores the default value 120.
maximum-prefix <*maximum-prefix*>[<*threshold*>] **no maximum-prefix <*maximum-prefix* >** **no maximum-prefix**	Configure the maximum of RIP route; the **no maximum-prefix <*maximum-prefix* >** **no maximum-prefix** command cancels the limit.
timers basic <*update*> <*invalid*> <*garbage*> **no timers basic**	Adjust the update, timeout and garbage collection time, the **no timers basic** command restores the default configuration.

recv-buffer-size ***<size>*** **no recv-buffer-size**	The command configures the UDP receiving buffer size of the RIP; the **no recv-buffer-size** command restores the system default values.

3. Configure RIP-I/RIP-II toggling

（1） Configure the RIP version to be used in all ports

Command	Explanation
RIP configuration mode	
version { 1 | 2 } **no version**	Configure the versions of all the RIP data packets transmitted/received by the Layer 3 switch port sending/receiving the **no version** command restores the default configuration, version 2.

（2） Configure the RIP version to send/receive in all ports.

（3） Configure whether to enable RIP packets sending/receiving for ports

Command	Explanation
Interface configuration mode	
ip rip send version { 1 | 1-compatible | 2 } **no ip rip send version**	Sets the version of RIP packets to send on all ports; the **no ip rip send version** command set the version to the one configured by the version command.
ip rip receive version {1 | 2 | } **no ip rip receive version**	Sets the version of RIP packets to receive on all ports; the no action of this command set the version to the one configured by the version command.
ip rip receive-packet **no ip rip receive-packet**	Enables receiving RIP packets on the interface; the **no ip rip receive-packet** command close data receiving on this port.
ip rip send-packet **no ip rip send-packet**	Enables sending RIP packets on the interface; the **no ip rip send-packet** command disables sending RIP packets on the interface.

4. Delete the specified route in RIP route table

Command	Explanation
Admin Mode	
clear ip rip route {*<A.B.C.D/M>*|kernel|static|connected|rip|ospf|isis|bgp|all}	The command deletes a specified route from the RIP route table.

5. Configure the RIP routing aggregation

（1） **Configure IPv4 aggregation route globally**

Command	Explanation
Router Configuration Mode	
ip rip aggregate-address A.B.C.D/M **no ip rip aggregate-address A.B.C.D/M**	To configure or delete IPv4 aggregation route globally.

（2） **Configure IPv4 aggregation route on interface**

Command	Explanation
Interface Configuration Mode	
ip rip aggregate-address A.B.C.D/M **no ip rip aggregate-address A.B.C.D/M**	To configure or delete IPv4 aggregation route on interface.

（3） **Display IPv4 aggregation route information**

Command	Explanation
Admin Mode and Configuration Mode	
show ip rip aggregate	To display aggregation route information.

6. Configure redistribution of OSPF routing to RIP

(1) Enable Redistribution of OSPF routing to RIP

Command	Explanation
Router RIP Configuration Mode	
redistribute ospf [<*process-id*>] [metric <*value*>] [route-map <*word*>] **no redistribute ospf [<process-id>]**	To enable or disable the redistribution of OSPF routing to RIP.

(2) Display and debug the information about configuration of redistribution of OSPF routing to RIP

Command	Explanation
Admin Mode and Configuration Mode	
show ip rip redistribute	To display the information about configuration of redistribute from other routing.
Admin Mode	
debug rip redistribute message send **no debug rip redistribute message send** **debug rip redistribute route receive** **no debug rip redistribute route receive**	To enable or disable debugging messages sent by RIP for redistribution of OSPF routing. To enable or disable debugging messages received from NSM.

7. Configure VRF address family mode for RIP

Command	Explanation
Router RIP configuration mode	
address-family ipv4 vrf <*vrf-name*> **no address-family ipv4 vrf <*vrf-name*>**	The command configures a RIP address family on the VRF of the PE router; the no command deletes the configured address family.
Address family configuration mode	
exit-address-family	This command exits the address family mode.

4.3.3 RIP Examples

4.3.3.1 Typical RIP Examples

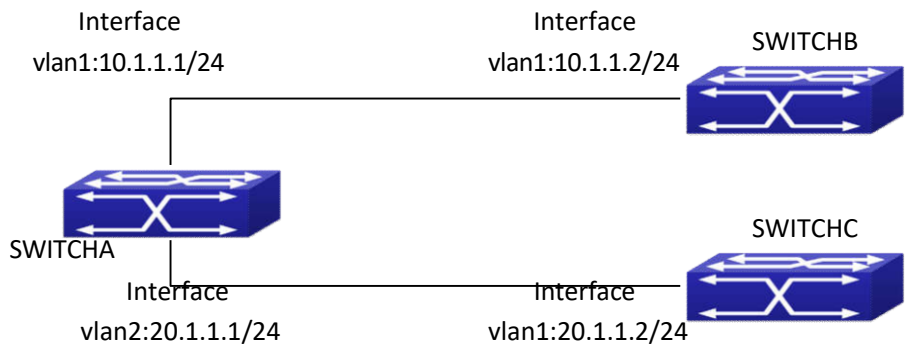

Figure 4-3 RIP example

In the figure shown above, a network consists of three Layer 3 switches, in which SwitchA connected with SwitchB and SwitchC, and RIP routing protocol is running in all of the three switches. SwitchA (interface vlan1：10.1.1.1,interface vlan2：20.1.1.1) exchanges Layer 3 switch update messages only with SwitchB (interface vlan1：10.1.1.2), but not with SwitchC (interface vlan 2: 20.1.1.2).

SwitchA, SwitchB, SwitchC configurations are as follows:

a) Layer 3 SwitchA：

Configure the IP address of interface vlan 1

SwitchA#config

SwitchA(config)# interface vlan 1

SwitchA(Config-if-Vlan1)# ip address 10.1.1.1 255.255.255.0

SwitchA(config-if-Vlan1)#

Configure the IP address of interface vlan 2

SwitchA(config)# vlan 2

SwitchA(Config-Vlan2)# switchport interface ethernet 1/0/2

Set the port Ethernet1/0/2 access vlan 2 successfully

SwitchA(Config-Vlan2)# exit

SwitchA(config)# interface vlan 2

SwitchA(Config-if-Vlan2)# ip address 20.1.1.1 255.255.255.0

Initiate RIP protocol and configure the RIP segments

SwitchA(config)#router rip

SwitchA(config-router)#network vlan 1

SwitchA(config-router)#network vlan 2

SwitchA(config-router)#exit

Configure that the interface vlan 2 do not transmit RIP messages to SwitchC

SwitchA(config)#router rip

SwitchA(config-router)#passive-interface vlan 2

SwitchA(config-router)#exit

SwitchA(config) #

b) Layer 3 SwitchB

Configure the IP address of interface vlan 1

SwitchB#config

SwitchB(config)# interface vlan 1

SwitchB(Config-if-Vlan1)# ip address 10.1.1.2 255.255.255.0
SwitchB(Config-if-Vlan1)exit
Initiate RIP protocol and configure the RIP segments
SwitchB(config)#router rip
SwitchB(config-router)#network vlan 1
SwitchB(config-router)#exit

c) Layer 3 SwitchC
SwitchC#config
SwitchC(config)# interface vlan 1
Configure the IP address of interface vlan 1
SwitchC(Config-if-Vlan1)# ip address 20.1.1.2 255.255.255.0
SwitchC(Config-if-Vlan1)#exit
Initiate RIP protocol and configure the RIP segments
SwitchC(config)#router rip
SwitchC(config-router)#network vlan 1
SwitchC(config-router)#exit

4.3.3.2 Typical Examples of RIP aggregation function

The application topology as follows：

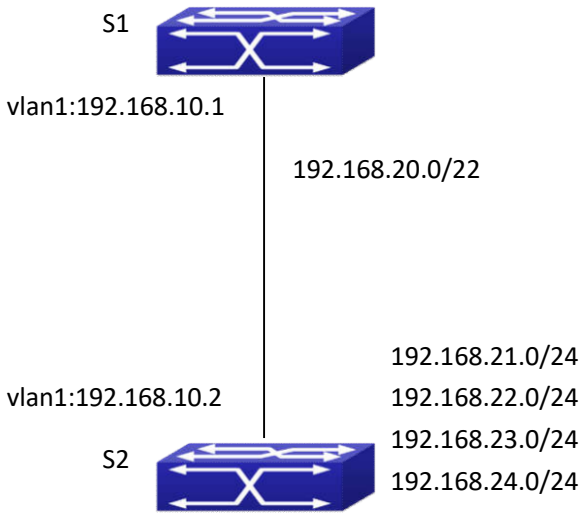

Figure 4-4 Typical application of RIP aggregation

As the above network topology, S2 is connected to S1 through interface vlan1, there are other 4 subnet routers of S2, which are 192.168.21.0/24, 192.168.22.0/24, 192.168.23.0/24, 192.168.24.0/24. S2 supports route aggregation, and to configure aggregation route 192.168.20.0/22 in interface vlan1 of S2, after that, sending router messages to S1 through vlan1, and put the four subnet routers aggregated to one router as 192.168.20.0/22, and send to S1, and not send subnet to neighbor. It can reduce the router table of S1, save the memory.

S1 configuration list:
S1(config)#router rip
S1(config-router) #network vlan 1

S2 configuration list:
S2(config)#router rip
S2(config-router) #network vlan 1
S2(config-router) #exit
S2(config)#in vlan 1
S2(Config-if-Vlan1)# ip rip agg 192.168.20.0/22

4.3.4 RIP Troubleshooting

The RIP protocol may not be working properly due to errors such as physical connection, configuration error when configuring and using the RIP protocol. So users should pay attention to following:

☞ First ensure the physic connection is correct

☞ Second, ensure the interface and chain protocol are UP (use **show interface** command)

☞ Then initiate the RIP protocol (use **router rip** command) and configure the segment (use **network** command) and set RIP protocol parameter on corresponding interfaces, such as the option between RIP-I and RIP-II

☞ After that, one feature of RIP protocol should be noticed ---the Layer 3 switch running RIP protocol sending route updating messages to all neighboring Layer 3 switches every 30 seconds. A Layer 3 switch is considered inaccessible if no route updating messages from the switch is received within 180 seconds, then the route to the switch will remains in the route table for 120 seconds before it is deleted. Therefore, if to delete a RIP route, this route item is assured to be deleted from route table after 300 seconds.

☞ When exchanging routing messages with CE using RIP protocol on the PE router, we should first create corresponding VPN routing/transmitting examples to associate with corresponding interfaces. Then enter the RIP address family mode configuring corresponding parameters. If the RIP routing problem remains unresolved, please use debug rip command to record the debug message in three minutes, and send them to our technical service center.

4.4 OSPF

4.4.1 Introduction to OSPF

OSPF is abbreviation for Open Shortest Path First. It is an interior dynamic routing protocol for autonomous system based on link-state. The protocol creates a link-state database by exchanging link-states among layer3 switches, and then uses the Shortest Path First algorithm to generate a route table basing on that database.

Autonomous system (AS) is a self-managed interconnected network. In large networks, such as the Internet, a giant interconnected network is broken down to autonomous systems. Big enterprise networks connecting to the Internet are independent AS, since the other hosts on the Internet are not managed by those AS and they don't share interior routing information with the layer3 switches on the Internet.

Each link-state Layer3 switch can provide information about the topology with its neighboring Layer3 switches.

- The network segment (link) connecting to the layer3 switch
- State of the connecting link

Link-state information is flooded throughout the network so that all Layer3 switches can get firsthand information. Link-state Layer3 switches will not broadcast all information contained in their route tables; instead, they only send changed link-state information. Link-state Layer3 switches establish neighborhood by sending "HELLO" to their neighbors, then link-state advertisements (LSA) will be sent among neighboring Layer3 switches. Neighboring Layer3 switch copy the LSA to their routing table and transfer the information to the rest part of the network. This process is referred to as "flooding". In this way, firsthand information is sent throughout the network to provide accurate map for creating and updating routes in the network. Link-state routing protocols use cost instead of hops to decide the route. Cost is assigned automatically or manually. According to the algorithm in link-state protocol, cost can be used to calculate the hop number for packets to pass, link bandwidth, and current load of the link. The administrator can even add weight for better assessment of the link-state.

1) When a link-state layer3 switch enters a link-state interconnected network, it sends a HELLO packet to get to know its neighbors and establish neighborhood.

2) The neighbors respond with information about the links they are connecting and the related costs.

3) The originate layer3 switch uses this information to build its own routing table

4) Then, as part of the regular update, layer3 switch send link-state advertisement (LSA) packets to its neighboring layer3 switches. The LSA include links and related costs of that layer3 switch.

5) Each neighboring layer3 switch copies the LSA packet and passes it to the next neighbor (i.e. flooding).

6) Since routing database is not recalculated before layer3 switch forwards LSA flooding, the converging time is greatly reduced.

One major advantage of link-state routing protocols is the fact that infinite counting is impossible, this is because of the way link-state routing protocols build up their routing table. The second advantage is that converging in a link-state interconnected network is very fast, once the routing topology changes, updates will be flooded throughout the network very soon. Those

advantages release some layer3 switch resources, as the process ability and bandwidth used by bad route information are minor.

The features of OSPF protocol include the following: OSPF supports networks of various scales, several hundreds of layer3 switches can be supported in an OSPF network. Routing topology changes can be quickly found and updating LSAs can be sent immediately, so that routes converge quickly. Link-state information is used in shortest path algorithm for route calculation, eliminating loop route. OSPF divides the autonomous system into areas, reducing database size, bandwidth occupation and calculation load. (According to the position of layer3 switches in the autonomous system, they can be grouped as internal area switches, area border switches, AS border switches and backbone switches). OSPF supports load balance and multiple routes to the same destination of equal costs. OSPF supports 4 level routing mechanisms (process routing according to the order of intra-area path, inter-area path, type 1 external path and type 2 external path). OSPF supports IP subnet and redistribution of routes from the other routing protocols, and interface-based packet verification. OSPF supports sending packets in multicast.

Each OSPF layer3 switch maintains a database describing the topology of the whole autonomous system. Each layer3 switch gathers the local status information, such as available interface, reachable neighbors, and sends link-state advertisement (sending out link-state information) to exchange link-state information with other OSPF layer3 switches to form a link-state database describing the whole autonomous system. Each layer3 switch builds a shortest path tree rooted by itself according to the link-state database, this tree provides the routes to all nodes in an autonomous system. If two or more layer3 switches exist (i.e. multi-access network), "designated layer3 switch" and "backup designated layer3 switch" will be selected. Designated layer3 switch is responsible for spreading link-state of the network. This concept helps reducing the traffic among the Layer3 switches in multi-access network.

OSPF protocol requires the autonomous system to be divided into areas. That is to divide the autonomous system into 0 area (backbone area) and non-0 areas. Routing information between areas are further abstracted and summarized to reduce the bandwidth required in the network. OSPF uses four different kinds of routes; they are intra-area route, inter-area route, type 1 external route and type 2 external route, in the order of highest priority to lowest. The route inside an area and between areas describes the internal network structure of an autonomous system, while external routes describe how to select the routing information to destination outside the autonomous system. The first type of exterior route corresponds to the information introduced by OSPF from the other interior routing protocols, the costs of those routes are comparable with the costs of OSPF routes; the second type of exterior route corresponds to the information introduced by OSPF from the other exterior routing protocols, but the costs of those routes are far greater than that of OSPF routes, so OSPF route cost is ignored when calculating route costs.

OSPF areas are centered with the Backbone area, identified as Area 0, all the other areas must be connected to Area 0 logically, and Area 0 must be continuous. For this reason, the concept of virtual link is introduced to the backbone area, so that physically separated areas still have logical connectivity to the backbone area. The configurations of all the layer3 switches in the same area must be the same.

In conclusion, LSA can only be transferred between neighboring Layer3 switches, OSPF protocol includes 5 types of LSA: router LSA, network LSA, network summary LSA to the other areas, ASBR summary LSA and AS external LSA. They can also be called type1 LSA, type2 LSA, type3 LSA, type4 LSA, and type5 LSA. Router LSA is generated by each layer3 switch inside an

OSPF area, and is sent to all the other neighboring layer3 switches in the same area; network LSA is generated by the designated layer3 switch in the OSPF area of multi-access network, and is sent to all other neighboring layer3 switches in this area. (In order to reduce traffic on layer3 switches in the multi-access network, “designated layer3 switch” and “backup designated layer3 switch” should be selected in the multi-access network, and the network link-state is broadcasted by the designated layer3 switch); network summary LSA is generated by border switches in an OSPF area , and is transferred among area border layer3 switches; AS external LSA is generated by layer3 switches on external border of AS, and is transferred throughout the AS.

As to autonomous systems mainly advertises exterior link-state, OSPF allow some areas to be configured as STUB areas to reduce the size of the topology database. Type4 LSA (ASBR summary LSA) and type5 LSA (AS external LSA) are not allowed to flood into/through STUB areas. STUB areas must use the default routes, the layer3 switches on STUB area edge advertise the default routes to STUB areas by type 3 summary LSA, those default routes only floods inside STUB area and will not get out of STUB area. Each STUB area has a corresponding default route, the route from a STUB area to AS exterior destination must rely on the default route of that area.

The following simply outlines the route calculation process of OSPF protocol:

1） Each OSPF-enabled layer3 switch maintains a database (LS database) describing the link-state of the topology structure of the whole autonomous system. Each layer3 switch generates a link-state advertisement according to its surrounding network topology structure (router LSA), and sends the LSA to other layer3 switches through link-state update (LSU) packets. Thus each layer3 switches receives LSAs from other layer3 switches, and all LSAs are combined to the link-state database.

2） Since a LSA is the description of the network topology structure around a layer3 switch, the LS database is the description of the network topology structure of the whole network. The layer3 switches can easily create a weighted vector map according to the LS database. Obviously, all layer3 switches in the same autonomous system will have the same network topology map.

3） Each layer3 switch uses the shortest path first (SPF) algorithm to calculate a tree of shortest path rooted by itself. The tree provides the route to all the nodes in the autonomous system, leaf nodes consist of the exterior route information. The exterior route can be marked by the layer3 switch broadcast it, so that additional information about the autonomous system can be recorded. As a result, the route table of each layer3 switch is different.

OSPF protocol is developed by the IETF; the OSPF v2 widely used now is fulfilled according to the content described in RFC2328.

4.4.2 OSPF Configuration Task List

The OSPF configuration may be different from the configuration procedure to switches of the other manufacturers. It is a two-step process:

1、 Enable OSPF in the Global Mode; 2、Configure OSPF area for the interfaces. The configuration task list is as follows:

1. Enable OSPF protocol (required)
   (1) Enable/disable OSPF protocol (required)
   (2) Configure the ID number of the layer3 switch running OSPF (optional)
   (3) Configure the network scope for running OSPF (optional)
   (4) Configure the area for the interface (required)
2. Configure OSPF protocol parameters (optional)
   (1) Configure OSPF packet sending mechanism parameters
      1) Configure OSPF packet verification
      2) Set the OSPF interface to receive only
      3) Configure the cost for sending packets from the interface
      4) Configure OSPF packet sending timer parameter (timer of broadcast interface sending HELLO packet to poll, timer of neighboring layer3 switch invalid timeout, timer of LSA transmission delay and timer of LSA retransmission.
   (2) Configure OSPF route introduction parameters
      1) Configure default parameters (default type, default tag value, default cost)
      2) Configure the routes of the other protocols to introduce to OSPF.
   (3) Configure OSPF importing the routes of other OSPF processes
      1) Enable the function of OSPF importing the routes of other OSPF processes
      2) Display relative information
      3) Debug
   (4) Configure other OSPF protocol parameters
      1) Configure OSPF routing protocol priority
      2) Configure cost for OSPF STUB area and default route
      3) Configure OSPF virtual link
      4) Configure the priority of the interface when electing designated layer3 switch (DR).
      5) Configure to keep a log for OSPF adjacency changes or not
      6) Filter the route obtained by OSPF
3. Disable OSPF protocol

1. Enable OSPF protocol

Basic configuration of OSPF routing protocol on switch is quite simple, usually only enabling OSPF and configuration of the OSPF area for the interface are required. The OSPF protocol parameters can use the default settings. If OSPF protocol parameters need to be modified, please refer to “2. Configure OSPF protocol parameters”.

Command	Explanation
Global Mode	
[no] router ospf [process <*id*>] [VRF Name]	Enables OSPF protocol; the **no** command disables OSPF protocol. (required)

OSPF Protocol Configuration Mode	
router-id <*router_id*> no **router-id**	Configures the ID number for the layer3 switch running OSPF; the **no router id** command cancels the ID number. The IP address of an interface is selected to be the layer3 switch ID. (optional)
[no] network {***<network> <mask>/ <network>/<prefix>***} **area <*area_id*>**	Configure certain segment to certain area, the **no network** {<network> ***<mask> / <network>/<prefix>}*** **area <area_*id*> *command*** cancels this configuration. (required)

2. Configure OSPF protocol parameters

《1》**Configure OSPF packet sending mechanism parameters**

1）Configure OSPF packet verification

2）Set the OSPF interface to receive only

3）Configure the cost for sending packets from the interface

Command	Explanation
Interface Configuration Mode	
ip ospf authentication { message-digest | null} **no ip ospf authentication**	Configures the authentication method by the interface to accept OSPF packets; the **no ip ospf authentication** command restores the default settings.
ip ospf [<*ip-address*>] authentication-key <*0 LINE | 7 WORD | LINE*> **no ip ospf [<*ip-address*>] authentication**	Specify the authentication key required in sending and receiving OSPF packet on the interface; the no command cancels the authentication key.
[no] passive-interface <*ifname*> [<*ip-address*>]	Sets an interface to receive only, the **no passive-interface <*ifname*>[<*ip-address*>]** command cancels this configuration.
ip ospf cost <*cost* > **no ip ospf cost**	Sets the cost for running OSPF on the interface; the **no ip ospf cost** command restores the default setting.

4) Configure OSPF packet sending timer parameter (timer of broadcast interface sending HELLO packet to poll, timer of neighboring layer3 switch invalid timeout, timer of LSA transmission delay and timer of LSA retransmission).

Command	Explanation
Interface Configuration Mode	
ip ospf hello-interval <*time*> **no ip ospf hello-interval**	Sets interval for sending HELLO packets; the **no ip ospf hello-interval** command restores the default setting.

ip ospf dead-interval <*time* > **no ip ospf dead-interval**	Sets the interval before regarding a neighbor layer3 switch invalid; the **no ip ospf dead-interval** command restores the default setting.
ip ospf transit-delay <*time*> **no ip ospf transit-delay**	Sets the delay time before sending link-state broadcast; the **no ip ospf transmit-delay** command restores the default setting.
ip ospf retransmit <*time*> **no ip ospf retransmit**	Sets the interval for retransmission of link-state advertisement among neighbor layer3 switches; the **no ip ospf retransmit** command restores the default setting.

（2）Configure OSPF route introduction parameters

Configure the routes of the other protocols to introduce to OSPF.

Command	Explanation
OSPF Protocol Configuration Mode	
redistribute { bgp | connected | static | rip | kernel} [metric-type { 1 | 2 }] [tag <*tag*>] [metric <*cost_value*>] [router-map <*WORD*>] **no redistribute { bgp | connected | static | rip | kernel }**	Distribute other protocols to find routing and static routings as external routing messages the **no redistribute {bgp | connected | static | rip | kernel}** command cancels the distributed external messages.

（3）Configure OSPF importing the routes of other OSPF processes

1） Enable the function of OSPF importing the routes of other OSPF processes

Command	Explanation
Router OSPF Mode	
redistribute ospf [<*process-id*>] [metric<*value*>] [metric-type {1|2}][route-map<*word*>] **no redistribute ospf [<*process-id*>] [metric<*value*>] [metric-type {1|2}][route-map<*word*>]**	Enable or disable the function of OSPF importing the routes of other OSPF processes.

2） Display relative information

Command	Explanation
Admin Mode or Configure Mode	
show ip ospf [<*process-id*>] redistribute	Display the configuration information of the OSPF process importing other outside routes.

3） Debug

Command	Explanation
Admin Mode	

debug ospf redistribute message send **no debug ospf redistribute message send** **debug ospf redistribute route receive** **no debug ospf redistribute route receive**	Enable or disable debugging of sending command from OSPF process redistributed to other OSPF process routing. Enable or disable debugging of received routing message from NSM for OSPF process.

《4》 **Configure other OSPF protocol parameters**

1）Configure how to calculate OSPF SPF algorithm time

2）Configure the LSA limit in the OSPF link state database

3）Configure various OSPF parameters

Command	Explanation
OSPF Protocol Configuration Mode	
timers spf <*interval*> **no timers spf**	Configure the SPF timer of OSPF; the **no timers spf** command restores the default settings.
overflow database {<*max-LSA*> [hard | soft] | external <*max-LSA*> <*recover time*>} **no overflow database [external <*max-LSA*> < *recover time* >]**	Configure the LSA limit in current OSPF process database; the **no overflow database [external < *max-LSA* > < *recover time* >]** command restores the default settings.
area <id> {authentication [message-digest] | default-cost <*cost*> | filter-list {access | prefix} <*WORD*> {in | out} | nssa [default-information-originate | no-redistribution | no-summary | translator-role] | range <*range*> | stub [no-summary] | virtual-link <*neighbor*>} **no area <id> {authentication | default-cost | filter-list {access | prefix} <*WORD*> {in | out} | nssa [default-information-originate | no-redistribution | no-summary | translator-role] | range <*range*> | stub [no-summary] | virtual-link <*neighbor*>}**	Configure the parameters in OSPF area (STUB area, NSSA area and virtual links); the **no area <id> {authentication | default-cost | filter-list {access | prefix} <*WORD*> {in | out} | nssa [default-information-originate | no-redistribution | no-summary | translator-role] | range <*range*> | stub [no-summary] | virtual-link <*neighbor*>}** command restores the default settings.

4）Configure the priority of the interface when electing designated layer3 switch (DR).

Command	Explanation
Interface Configuration Mode	
ip ospf priority <*priority*> **no ip ospf priority**	Sets the priority of the interface in "designated layer3 switch" election; the **no ip ospf priority** command restores the default setting.

5）Configure to keep a log for OSPF adjacency changes or not

Command	Explanation
OSPF Protocol Configuration Mode	

log-adjacency-changes detail **no log-adjacency-changes detail**	Configure to keep a log for OSPF adjacency changes or not.

6) Filter the route obtained by OSPF

Command	Explanation
OSPF Protocol Configuration Mode	
filter-policy <access-list-name> **no filter-policy**	Use access list to filter the route obtained by OSPF, the no command cancels the route filtering.

3. Disable OSPF protocol

Command	Explanation
Global Mode	
no router ospf [process <*id*>] Disables OS	F routing protocol.

4.4.3 OSPF Examples

4.4.3.1 Configuration Example of OSPF

Scenario 1: OSPF autonomous system.

This scenario takes an OSPF autonomous system consists of five switch for example.

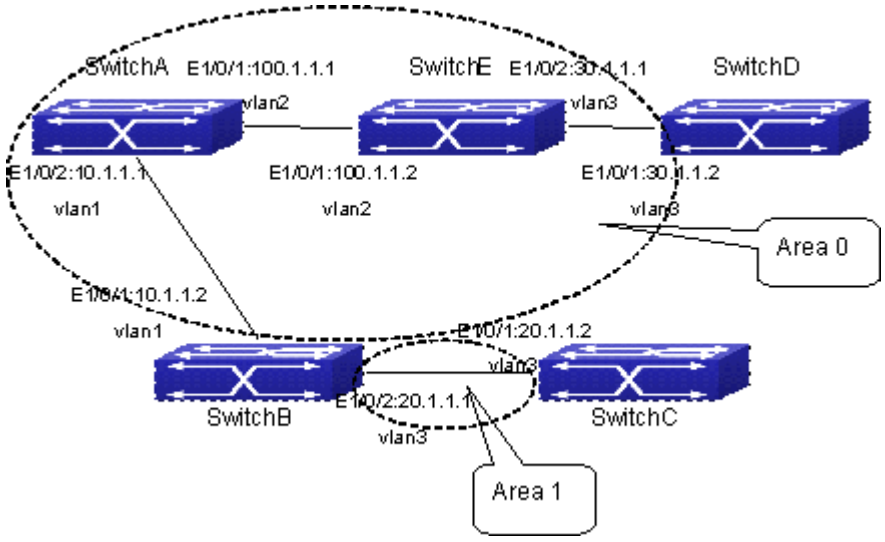

Figure 4-5 Network topology of OSPF autonomous system

The configuration for layer3 Switch1 and Switch5 is shown below:

Layer 3 Switch1

Configuration of the IP address for interface vlan1

Switch1#config

Switch1(config)# interface vlan 1

Switch1(config-if-vlan1)# ip address 10.1.1.1 255.255.255.0

Switch1(config-if-vlan1)#exit

Configuration of the IP address for interface vlan2

Configure the IP address of interface vlan2

Switch1(config)# interface vlan 2
Switch1(config-if-vlan2)# ip address 100.1.1.1 255.255.255.0
Switch1 (config-if-vlan2)#exit
Enable OSPF protocol, configure the area number for interface vlan1 and vlan2.
Switch1(config)#router ospf
Switch1(config-router)#network 10.1.1.0/24 area 0
Switch1(config-router)#network 100.1.1.0/24 area 0
Switch1(config-router)#exit
Switch1(config)#exit
Switch1#
Layer 3 Switch2:
Configure the IP address for interface vlan1 and vlan2.
Switch2#config
Switch2(config)# interface vlan 1
Switch2(config-if-vlan1)# ip address 10.1.1.2 255.255.255.0
Switch2(config-if-vlan1)#no shutdown
Switch2(config-if-vlan1)#exit
Switch2(config)# interface vlan 3
Switch2(config-if-vlan3)# ip address 20.1.1.1 255.255.255.0
Switch2(config-if-vlan3)#no shutdown
Switch2(config-if-vlan3)#exit
Enable OSPF protocol, configure the OSPF area interfaces vlan1 and vlan3 in
Switch2(config)#router ospf
Switch2(config-router)# network 10.1.1.0/24 area 0
Switch2(config-router)# network 20.1.1.0/24 area 1
Switch2(config-router)#exit
Switch2(config)#exit
Switch2#
Layer 3 Switch3:
Configuration of the IP address for interface vlan3.
Switch3#config
Switch3(config)# interface vlan 3
Switch3(config-if-vlan1)# ip address 20.1.1.2 255.255.255.0
Switch3(config-if-vlan3)#no shutdown
Switch3(config-if-vlan3)#exit
Initiate the OSPF protocol, configure the OSPF area to which interface vlan3 belongs
Switch3(config)#router ospf
Switch3(config-router)# network 20.1.1.0/24 area 1
Switch3(config-router)#exit
Switch3(config)#exit
Switch3#
Layer 3 Switch4:
Configuration of the IP address for interface vlan3
Switch4#config

Switch4(config)# interface vlan 3
Switch4(config-if-vlan3)# ip address30.1.1.2 255.255.255.0
Switch4(config-if-vlan3)#no shutdown
Switch4(config-if-vlan3)#exit
Enable OSPF protocol, configure the OSPF area interfaces vlan3 resides in.
Switch4(config)#router ospf
Switch4(config-router)# network 30.1.1.0/24 area 0
Switch4(config-router)#exit
Switch4(config)#exit
Switch4#
Layer 3 Switch5:
Configuration of the IP address for interface vlan2
Switch5#config
Switch5(config)# interface vlan 2
Switch5(config-if-vlan2)# ip address 100.1.1.2 255.255.255.0
Switch5(config-if-vlan2)#no shutdown
Switch5(config-if-vlan2)#exit
Configuration of the IP address for interface vlan3
Switch5(config)# interface vlan 3
Switch5(config-if-vlan3)# ip address 30.1.1.1 255.255.255.0
Switch5(config-if-vlan3)#no shutdown
Switch5(config-if-vlan3)#exit
Enable OSPF protocol, configure the number of the area in which interface vlan2 and vlan3 reside in.
Switch5(config)#router ospf
Switch5(config-router)# network 30.1.1.0/24 area 0
Switch5(config-router)# network 100.1.1.0/24 area 0
Switch5(config-router)#exit
Switch5(config)#exit
Switch5#

Scenario 2: Typical OSPF protocol complex topology.

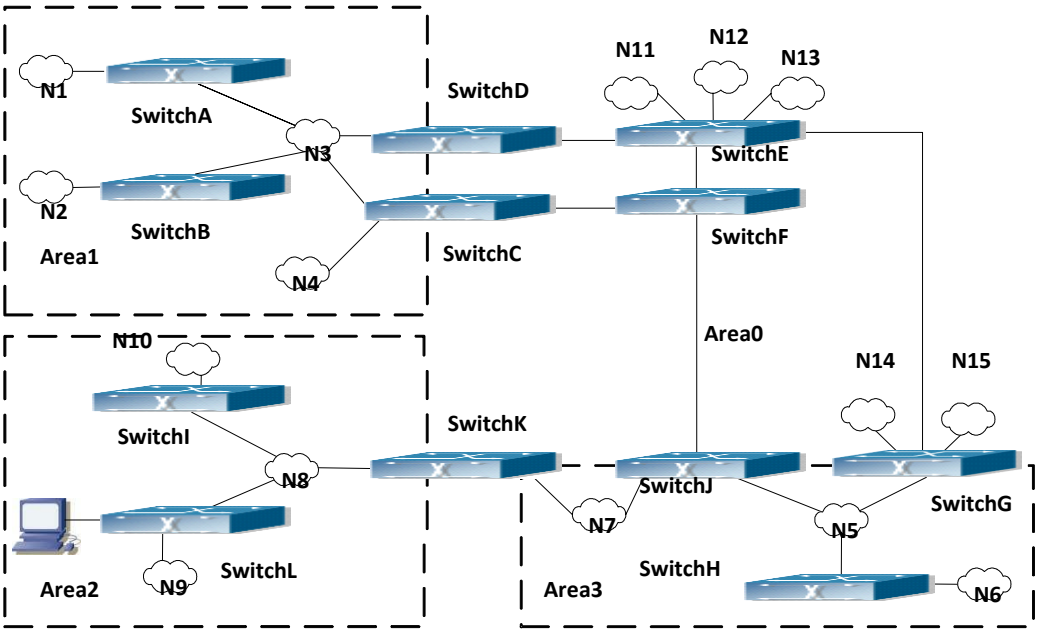

Figure 4-6 Typical complex OSPF autonomous system

This scenario is a typical complex OSPF autonomous system network topology. Area1 include network N1-N4 and layer3 SwitchA-SwitchD, area2 include network N8-N10, host H1 and layer3 SwitchH, area3 include N5-N7 and layer3 SwitchF, SwitchG SwitchA0 and Switch11, and network N8-N10 share a summary route with host H1(i.e. area3 is defined as a STUB area). Layer3 SwitchA, SwitchB, SwitchD, SwitchE, SwitchG, SwitchH, Switch12 are in-area layer3 switches, SwitchC, SwitchD, SwitchF, Switch10 and Switch11 are edge layer3 switches of the area, SwitchD and SwitchF are edge layer3 switches of the autonomous system.

To area1, layer3 switches SwitchA and SwitchB are both in-area switches, area edge switches SwitchC and SwitchD are responsible for reporting distance cost to all destination outside the area, while they are also responsible for reporting the position of the AS edge layer3 switches SwitchD and SwitchF, AS exterior link-state advertisement from SwitchD and SwitchF are flooded throughout the whole autonomous system. When ASE LSA floods in area 1, those LSAs are included in the area 1 database to get the routes to network N11 and N15.

In addition, layer3 SwitchC and SwitchD must summary the topology of area 1 to the backbone area (area 0, all non-0 areas must be connected via area 0, direct connections are not allowed), and advertise the networks in area 1 (N1-N4) and the costs from SwitchC and SwitchD to those networks. As the backbone area is required to keep connected, there must be a virtual link between backbone layer3 Switch10 and Switch11. The area edge layer3 switches exchange summary information via the backbone layer3 switch, each area edge layer3 switch listens to the summary information from the other edge layer3 switches.

Virtual link can not only maintain the connectivity of the backbone area, but also strengthen the backbone area. For example, if the connection between backbone layer3 SwitchG and Switch10 is cut down, the backbone area will become incontinuous. The backbone area can become more robust by establishing a virtual link between backbone layer3 switches SwitchF and Switch10. In addition, the virtual link between SwitchF and Switch10 provide a short path from

area 3 to layer3 SwitchF.

Take area 1 as an example. Assume the IP address of layer3 SwitchA is 10.1.1.1, IP address of layer3 SwitchB interface VLAN2 is 10.1.1.2, IP address of layer3 SwitchC interface VLAN2 is 10.1.1.3, IP address of layer3 SwitchD interface VLAN2 is 10.1.1.4. SwitchA is connecting to network N1 through Ethernet interface VLAN1 (IP address 20.1.1.1); SwitchB is connecting to network N2 through Ethernet interface VLAN1 (IP address 20.1.2.1); SwitchC is connecting to network N4 through Ethernet interface VLAN3 (IP address 20.1.3.1). All the three addresses belong to area 1. SwitchC is connecting to layer3 SwitchE through Ethernet interface VLAN1 (IP address 10.1.5.1); SwitchD is connecting to layer3 SwitchD through Ethernet interface VLAN1 (IP address 10.1.6.1); both two addresses belong to area 1. Simple authentication is implemented among layer3 switches in area1, edge layer3 switches of area 1 authenticate with the area 0 backbone layer3 switches by MD5 authentication.

The followings are just configurations for all layer3 switches in area 1, configurations for layer3 switches of the other areas are omitted. The following are the configurations of SwitchA SwitchB.SwitchC and SwitchD:

1) SwitchA:

Configure IP address for interface vlan2

```
SwitchA#config
SwitchA(config)# interface vlan 2
SwitchA(config-If-Vlan2)# ip address 10.1.1.1 255.255.255.0
SwitchA(config-If-Vlan2)#exit
```

Enable OSPF protocol, configure the area number for interface vlan2.

```
SwitchA(config)#router ospf
SwitchA(config-router)#network 10.1.1.0/24 area 1
SwitchA(config-router)#exit
```

Configure simple key authentication.

```
SwitchA(config)#interface vlan 2
SwitchA(config-If-Vlan2)#ip ospf authentication
SwitchA(config-If-Vlan2)#ip ospf authentication-key test
SwitchA(config-If-Vlan2)exit
```

Configure IP address and area number for interface vlan1.

```
SwitchA(config)# interface vlan 1
SwitchA(config-If-Vlan1)#ip address 20.1.1.1 255.255.255.0
SwitchA(config-If-Vlan1)#exit
SwitchA(config)#router ospf
SwitchA(config-router)#network 20.1.1.0/24 area 1
SwitchA(config-router)#exit
```

2) SwitchB:

Configure IP address for interface vlan2

```
SwitchB#config
SwitchB(config)# interface vlan 2
SwitchB(config-If-Vlan2)# ip address 10.1.1.2 255.255.255.0
SwitchB(config-If-Vlan2)#exit
```

Enable OSPF protocol, configure the area number for interface vlan2.

SwitchB(config)#router ospf
SwitchB(config-router)#network 10.1.1.0/24 area 1
SwitchB(config-router)#exit
SwitchB(config)#interface vlan 2
Configure simple key authentication.
SwitchB(config)#interface vlan 2
SwitchB(config-If-Vlan2)#ip ospf authentication
SwitchB(config-If-Vlan2)#ip ospf authentication-key test
SwitchB(config-If-Vlan2)#exit
Configure IP address and area number for interface vlan1.
SwitchB(config)# interface vlan 1
SwitchB(config-If-Vlan1)#ip address 20.1.2.1 255.255.255.0
SwitchB(config-If-Vlan1)#exit
SwitchB(config)#router ospf
SwitchB(config-router)#network 20.1.2.0/24 area 1
SwitchB(config-router)#exit
SwitchB(config)#exit
3)SwitchC:
Configure IP address for interface vlan2
SwitchC#config
SwitchC(config)# interface vlan 2
SwitchC(config-If-Vlan2)# ip address 10.1.1.3 255.255.255.0
SwitchC(config-If-Vlan2)#exit
Enable OSPF protocol, configure the area number for interface vlan2
SwitchC(config)#router ospf
SwitchC(config-router)#network 10.1.1.0/24 area 1
SwitchC(config-router)#exit
Configure simple key authentication
SwitchC(config)#interface vlan 2
SwitchC(config-If-Vlan2)#ip ospf authentication
SwitchC(config-If-Vlan2)#ip ospf authentication-key test
SwitchC(config-If-Vlan2)#exit
Configure IP address and area number for interface vlan3
SwitchC(config)# interface vlan 3
SwitchC(config-If-Vlan3)#ip address 20.1.3.1 255.255.255.0
SwitchC(config-If-Vlan3)#exit
SwitchC(config)#router ospf
SwitchC(config-router)#network 20.1.3.0/24 area 1
SwitchC(config-router)#exit
Configure IP address and area number for interface vlan 1
SwitchC(config)# interface vlan 1
SwitchC(config-If-Vlan1)#ip address 10.1.5.1 255.255.255.0
SwitchC(config-If-Vlan1)#exit
SwitchC(config)#router ospf

SwitchC(config-router)#network 10.1.5.0/24 area 0
SwitchC(config-router)#exit
Configure MD5 key authentication.
SwitchC(config)#interface vlan 1
SwitchC (config-If-Vlan1)#ip ospf authentication message-digest
SwitchC (config-If-Vlan1)#ip ospf authentication-key test
SwitchC (config-If-Vlan1)#exit
SwitchC(config)#exit
SwitchC#
4)SwitchD:
Configure IP address for interface vlan2
SwitchD#config
SwitchD(config)# interface vlan 2
SwitchD(config-If-Vlan2)# ip address 10.1.1.4 255.255.255.0
SwitchD(config-If-Vlan2)#exit
Enable OSPF protocol, configure the area number for interface vlan2.
SwitchD(config)#router ospf
SwitchD(config-router)#network 10.1.1.0/24 area 1
SwitchD(config-router)#exit
Configure simple key authentication.
SwitchD(config)#interface vlan 2
SwitchD(config-If-Vlan2)#ip ospf authentication
SwitchD(config-If-Vlan2)#ip ospf authentication-key test
SwitchD(config-If-Vlan2)#exit
Configure the IP address and the area number for the interface vlan 1
SwitchD(config)# interface vlan 1
SwitchD(config-If-Vlan1)# ip address 10.1.6.1 255.255.255.0
SwitchD(config-If-Vlan1)exit
SwitchD(config)#router ospf
SwitchD(config-router)#network 10.1.6.0/24 area 0
SwitchD(config-router)#exit
Configure MD5 key authentication
SwitchD(config)#interface vlan 1
SwitchD(config-If-Vlan1)#ip ospf authentication message-digest
SwitchD(config-If-Vlan1)#ip ospf authentication-key test
SwitchD(config-If-Vlan1)exit
SwitchD(config)#exit
SwitchD#

Scenario 3: The function of OSPF importing the routers of other OSPF processes

As shown in the following graph, a switch running the OSPF routing protocol connects two networks: network A and network B. Because of some reason, it is required that network A should be able to learn the routers of network B, but network B should not be able to learn the routers of network A. According to that, two OSPF processes can be started respectively on

interface vlan 1 and interface vlan 2. the OSPF process which interface vlan 1 belongs to is configured to import the routers of the OSPF process which interface vlan 2 belongs to, while the OSPF process which interface vlan 2 belongs to should not be configured to import the routers of the OSPF process which interface vlan 1 belongs to.

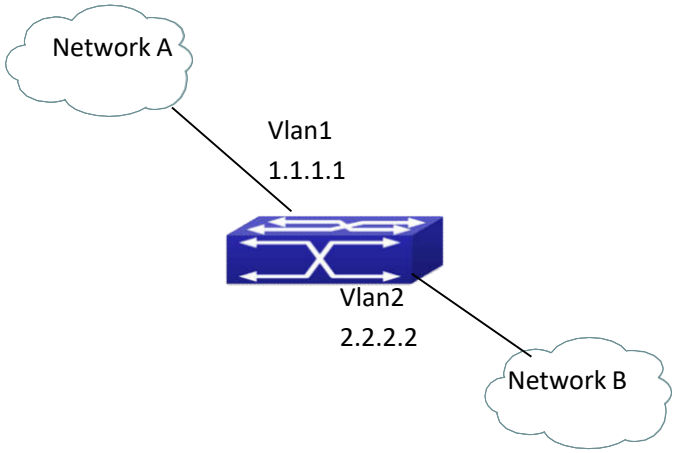

Figure 4-7 Function of OSPF importing the routers of other OSPF processes example

We can configure as follows:

```
Switch(config)#interface vlan 1
Switch(Config-if-Vlan1)#ip address 1.1.1.1 255.255.255.0
Switch(Config-if-Vlan1)#exit
Switch(config)#interface vlan 2
Switch(Config-if-Vlan2)#ip address 2.2.2.2 255.255.255.0
Switch(Config-if-Vlan2)#exit
Switch(config)#router ospf 10
Switch(config-router)#network 2.2.2.0/24 area 1
Switch(config-router)#exit
Switch(config)#router ospf 20
Switch(config-router)#network 1.1.1.0/24 area 1
Switch(config-router)#redistribute ospf 10
Switch(config-router)#exit
```

4.4.3.2 Configuration Examples of OSPF VPN

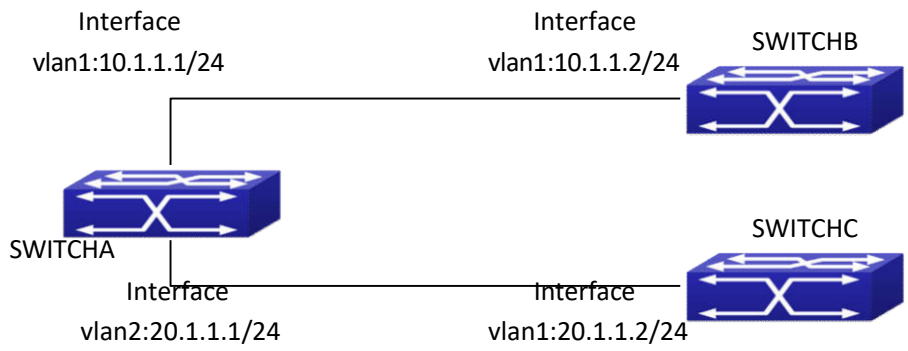

Figure 4-8 OSPF VPN Example

The above figure shows that a network consists of three Layer 3 switches in which the switchA as PE, SwitchB and SwitchC as CE1 and CE2. The PE is connected to CE1 and CE2 through vlan1 and vlan2. The routing messages are exchanged between PE and CE through OSPF protocol.

a) SwitchA, the Layer 3 switch as PE

Configure VPN route/transmitting examples vpnb and vpnc

SwitchA#config

SwitchA(config)#ip vrf vpnb

SwitchA(config-vrf)#

SwitchA(config-vrf)#exit

SwitchA#(config)

SwitchA(config)#ip vrf vpnc

SwitchA(config-vrf)#

SwitchA(config-vrf)#exit

Associate the vlan 1 and vlan 2 respectively with vpnb and vpnc while configuring IP address

SwitchA(config)#in vlan1

SwitchA(config-if-Vlan1)#ip vrf forwarding vpnb

SwitchA(config-if-Vlan1)#ip address 10.1.1.1 255.255.255.0

SwitchA(config-if-Vlan1)#exit

SwitchA(config)#in vlan2

SwitchA(config-if-Vlan2)#ip vrf forwarding vpnc

SwitchA(config-if-Vlan2)#ip address 20.1.1.1 255.255.255.0

SwitchA(config-if-Vlan2)#exit

Configure OSPF examples associated with vpnb and vpnc respectively

SwitchA(config)#

SwitchA(config)#router ospf 100 vpnb

SwitchA(config-router)#network 10.1.1.0/24 area 0

SwitchA(config-router)#redistribute bgp

SwitchA(config-router)#exit

SwitchA(config)#router ospf 200 vpnc

SwitchA(config-router)#network 20.1.1.0/24 area 0

SwitchA(config-router)#redistribute bgp

b) The Layer 3 SwitchB of CE1：

Configure the IP address of Ethernet E 1/0/2

SwitchB#config
SwitchB(config)# interface Vlan1
SwitchB(config-if-vlan1)# ip address 10.1.1.2 255.255.255.0
SwitchB (config-if-vlan1)exit
Enable OSPF protocol and configuring OSPF segments
SwitchB(config)#router ospf
SwitchB(config-router-rip)#network 10.1.1.0/24 area 0
SwitchB(config-router-rip)#exit

c) The Layer 3 SwitchC of CE2
Configure the IP address of Ethernet E 1/0/2
SwitchC#config
SwitchC(config)# interface Vlan1
SwitchC(config-if-vlan1)# ip address 20.1.1.2 255.255.255.0
SwitchC(config-if-vlan1)#exit
Initiate OSPF protocol and configuring OSPF segments
SwitchC(config)#router ospf
SwitchC(config-router)#network 20.1.1.0/24 area 0
SwitchC(config-router)#exit

4.4.4 OSPF Troubleshooting

The OSPF protocol may not be working properly due to errors such as physic connection, configuration error when configuring and using the OSPF protocol. So users should pay attention to following:

- ☞ First ensure the physic connection is correct
- ☞ Second, ensure the interface and link protocol are UP (use **show interface** command)
- ☞ Configure different IP address from different segment on each interface
- ☞ Then initiate OSPF protocol (use **router-ospf** command) and configure the OSPF area on corresponding interface
- ☞ After that, a OSPF protocol feature should be checked---the OSPF backbone area should be continuous and apply virtual link to ensure it is continuous. if not; all non 0 areas should only be connected to other non 0 area through 0 area; a border Layer 3 switch means that one part of the interfaces of this switch belongs to 0 area, the other part belongs to non 0 area; Layer 3 switch DR should be specified for multi-access network such as broadcast network.
- ☞ DHCP Relay doesn’t work between VRF.

4.5 BGP

4.5.1 Introduction to BGP

BGP stands for a Border Gateway Protocol. It's a dynamic routing protocol inter-autonomous system. Its basic function is automatically exchanging routing information without loops. By exchanging routing reachable information with autonomous number of AS sequence attributes, BGP could create autonomous topological map to eliminate routing loop and implement policies configured by users. Generally, the switches in an AS may use several IGPs (Interior Gateway Protocol) in order to exchange routing information in the AS, such as RIP and OSPF which are IGPs; and exchange information among ASes with EGP (Exterior Gateway Protocol). For example, BGP is one kind of EGP. The AS is usually established on a single administrative department. BGP is often used on the switches among ISPs or the departments of Multi-national Corporation.

BGP has been used since1989, its earliest three versions are RFC1105 (BGP-1), RFC1163 (BGP-2) and RFC1267 (BGP-3).Currently, the most popular one is RFC1771 (BGP-4). The switch supports BGP-4.

1、 Characteristics of BGP-4

BGP-4 is suitable for the distributed structure and supports Classless InterDomain Routing (CIDR). BGP-4 is becoming the virtual exterior routing protocol standard used for the global Internet. The features of BGP-4 are as follows.

- BGP is an exterior routing protocol, unlike interior routing protocol, such as OSPF and RIP, BGP can't discovery and calculate routes, but it can control the transmission of routes and select the best route.
- By carrying AS routing information in the updating route, the problem of Routing Loops can be resolved
- BGP uses TCP on port 179 as its transport protocol, this could enhance the reliability of the protocol.
- BGP-4 supports CIDR (Classless InterDomain Routing), which is an important improvement to BGP-3. CIDR has a brand new way to look on IP address; it doesn't distinguish class A , Class B and class C network. For instance, an illegal class C address 192.213.0.0 255.255.0.0 can be represented as 192.213.0.0/16 by CIDR which is a legal super network. /16 represents that the network number is formed by 16 bits from the beginning left of the address. The introduction of CIDR abbreviates the route aggregation. The route aggregation is the process of combining several different routes. So notifying several routes can be changed to notify only one route which decreases the route table.
- When updating route, BGP send only incremental route. The bandwidth occupied by BGP transmission is reduced greatly and it is suitable for the mass routing information transmitted on the internet
- For political and economical reasons, each AS expects to filter and control the route,

BGP-4 provides abundant route policies which make BGP-4 more extendable to encourage the internet development.

2、 The Overview of BGP-4 operation

Unlike RIP and OSPF protocols, BGP protocol is connection oriented. BGP switches must establish connection to exchange routing information. The operation of BGP protocol is driven by messages and the messages can be divided into four kinds:

Open message It's the first message which is sent after a TCP connection is established. It is used to create BGP connecting relation among BGP peers. Some parameters in Open Message are used to negotiate if a connection could be established among BGP peers.

Keepalive Message it's the message to check connection availability. It's usually sent periodically to keep BGP connection. If this message or Update message is not received within holdtime time, BGP connection is closed.

Update Message----- it's the most important message in the BGP system. It's used to exchange routing information among peers. The switches exchange not only updated routing information, but also unavailable or canceled routing information. It consists of three parts: unreachable route, NLRI (Network LayerReachability Information) and Path Attributes.

Notification Message------it's the mistake notification message. When a BGP speaker receives this message, it shutdowns the BGP connections with its neighbors

BGP-4 is connection oriented. BGP acts as higher protocol and runs on the particular equipments. When detecting a neighbor, a TCP session is established and maintained. Then the exchanging and synchronization of the route table will be carried out. By sending the whole BGP route table the routing information is exchanged only when the system initiates. After that, the routing information is exchanged only when the updated routing information is available. Only incremental update message is exchanged. BGP-4 maintains links and sessions periodically through keep alive message. That is sending and receiving keep alive message periodically to check if the connections are normal.

The switches that participate the BGP session are called BGP speaker. It continuously receives or generates new routing information and advertises it to other BGP speakers. When a BGP speaker receives a new routing notification from other AS, if this route is better than the presently known route or there is no acceptable route, it sends this route to all the other BGP speakers of the AS. A BGP speaker calls other speakers that exchange route information with it as neighbors or peers. Several relevant neighbors can constitute a peer group. BGP operates on the switches in the following two manners:

- IBGP：Internal BGP
- EBGP：External BGP

When BGP runs in the same AS, it's called IBGP. When in the different AS, it's called EBGP. Generally, the outer neighbors are connected physically and the inner neighbors can be in any place of the AS. The difference is finally shown in the dealing manner of BGP to routing information. The equipments may check the AS numbers of the Open Message from neighbors to decide treating the neighbor switches as the exterior neighbor or as the interior neighbor.

IBGP are used in the AS. It sends message to all the BGP neighbors in the AS. IBGP exchanges AS routing information in a big organization. Attention, the switches in the AS needn't be connected physically. Only if the switches are in the same AS, they can be neighbors each other. Because BGP can't detect route, the route tables of other inner route protocols (such as static

route, direct route, OSPF and RIP) need contain neighbor IP addresses and these routes are used to exchange information among BGPs. In order to avoid routing loops, when a BGP speaker receives a route notification from inner neighbor, it would not notify this route to other inner neighbors.

EBGP is used among the AS, and it transmits routing information to the BGP neighbors of outer ASes. EBGP need physical connection and share the same medium. Because EBGP need physical connection, the boundary equipments between two AS are usually running EBGP. When a BGP speaker receives routing information from outer neighbors, it notifies these routes to other inner neighbors.

3、 Route attribute

BGP-4 can share and query inner IP route table through relevant mechanisms, but it has its own route table. In the BGP route table, each route has a network number, AS listing information (also called AS path) that it passed and some routing attributes (such as origin). The routing attribute that BGP-4 used is very complex, this attribute can be used as metrics to select path.

4、 Route-selecting policy of BGP

When receiving BGP notification about a same route from several neighbors, selecting the best route need to be take into account after routing filtering. This process is called BGP route selecting process. BGP route selecting process will start only when the following conditions are fulfilled:

- The switch's route must be next hop reachable. That is in the route table there is the route that can reach the next hop.
- BGP must be synchronized with IGP (unless asynchronism is configured; only restricted to IBGP)

BGP route selecting process is based on the BGP attribute. When there are several routes that indicate the same destination, BGP need select the best route to the destination. The decision-making process is as the following:

1. Select the route with the most weight first;
2. If the weights are the same, select the route with the most local preference;
3. If the local preferences are the same, select the route generated by local switch.
4. If the local preferences are the same and there is no route generated by local switch, select the route with the shortest AS path;
5. If the AS paths are the same, select the route with the lowest "origin" type (IGP<EGP<INCOMPLETE);
6. If the "origin" types are the same, select the route with the lowest MED attribute. Unless activating command "bgp always-compare-med", this comparison is only available among the routes from the same neighbor AS.
7. If the MED attributes are the same, EBGP is preferable to outer confederation and outer confederation is preferable to IBGP.
8. If it's still the same by now, BGP router ID （router ID） is used to break the balance. The best route is the one from the least router ID.
9. If it's still the same by now, BGP router ID （router ID） is used to break the balance. The best route is the one from the least router ID.

4.5.2 BGP Configuration Task List

The BGP configuration tasks include basic and advanced tasks. Basic BGP configuration tasks include the following:

1. Enable BGP Routing (required)
2. Configure BGP Neighbors (required)
3. Administrate the change of routing policy
4. Configure BGP Weights
5. Configure BGP Route Filtering policy basing on Neighbors
6. Configure Next-Hop of BGP
7. Configure Multi-Hop of EGBP
8. Configure BGP Session Identifier
9. Configure BGP Version

Advanced BGP configuration tasks include the following:

1. Use Route Maps to Modify Route
2. Configure Route Aggregation
3. Configure BGP Community Filtering
4. Configure BGP Confederation
5. Configure a Route Reflector
6. Configure Peer Groups
7. Configure Neighbors and Peer Groups' Parameters
8. Adjust BGP Timers
9. Adjust BGP Announcement Interval
10. Configure the default Local Priority
11. Allow to Transfer Default Route
12. Configure BGP's MED Value
13. Configure BGP Routing Redistribution
14. Configure BGP Route Dampening
15. Configure BGP capability Negotiation
16. Configure Routing Server
17. Configure Path-Selected Rule
18. Configure redistribution of OSPF routing to BGP
    (1) Enable redistribution of OSPF routing to BGP
    (2) Display and debug the information about configuration of redistribution of OSPF routing to BGP

Ⅰ. Basic BGP configuration tasks

1. Enable BGP Routing

Command	Explanation
Global mode	
router bgp <*as-id*> **no router bgp <*as-id*>**	Enable BGP, the **no router bgp <*as-id*>** command disables BGP process.
BGP protocol mode	
bgp *asnotation asdot* ***no* bgp asnotation asdot**	Show AS number and match the regular expression **with ASDOT method.** ***The no*** command cancels this method.
network <ip-address/M> **no network <ip-address/M>**	Set the network that BGP will announce, the no network <ip-address/M> command cancels the network that will be announced.
address-family ipv4 {unicast| multicast|vrf <*vrf-nam*>} **no address-family ipv4 {unicast| multicast|vrf <*vrf-nam*>}**	Create IPv4 for BGP protocol and enter BGP-VPN view. Any IPv4 is not created by default.

2. Configure BGP Neighbors

Command	Explanation
Router configuration mode	
neighbor {<*ip-address*>|<*TAG*>} remote-as <*as-id*> **no neighbor {<*ip-address*>|<*TAG*>} [remote-as <*as-id*>]**	Specify a BGP neighbor, the **no neighbor {<*ip-address*>|<*TAG*>} [remote-as <*as-id*>]** command deletes the neighbor.

3. Administrate the change of routing policy

（1） Configure hard reconfiguration.

Command	Explanation
Admin Mode	
clear ip bgp {<*>|<*as-id*>| external|peer-group <*NAME*>|<*ip-address*>}	Configure hard reconfiguration.

（2） Configure outbound soft reconfiguration.

Command	Explanation
Admin Mode	
clear ip bgp {<*>|<*as-id*>| external|peer-group <NAME>|<*ip-address*>} soft out	Configure outbound soft reconfiguration.

（3） Configure inbound soft reconfiguration.

Command	Explanation
BGP configuration mode	

neighbor { *<ip-address>* | *<TAG>* } soft-reconfiguration inbound **no neighbor { *<ip-address>* | *<TAG>* } soft-reconfiguration inbound**	This command can store routing information from neighbors and peers; the **no neighbor { *<ip-address>* | *<TAG>* } soft-reconfiguration inbound** command cancels the storage of routing information.
Admin Mode	
clear ip bgp {*<*>*|*<as-id>*| external|peer-group *<NAME>*|*<ip-address>*} soft in	Configure BGP inbound soft reconfiguration.

4. Configure BGP Weights

Command	Explanation
BGP configuration mode	
neighbor { *<ip-address>* | *<TAG>* } weight *<weight>* **no neighbor { <ip-address> | <TAG> }**	Configure BGP neighbor weights; the **no neighbor { <ip-address> | <TAG> }** command recovers default weights.

5. Configure BGP Route Filtering policy based on neighbor

Command	Explanation
BGP configuration mode	
neighbor {*<ip-address>*|*<TAG>*} distribute-list {*<1-199>*|*<1300-2699>*|*<WORD>*} {in|out} **no neighbor {*<ip-address>*|*<TAG>*} distribute-list {*<1-199>*|*<1300-2699>*|*<WORD>*} {in|out}**	Filter neighbor routing updating information. The **no neighbor {*<ip-address>* | *<TAG>*} distribute-list {*<1-199>*|*<1300-2699>*| *<WORD>*} {in|out}** command cancels routing filter.

6. Configure Next-Hop

1) Set Next-Hop as the switch's address

Command	Explanation
BGP configuration mode	
neighbor { *<ip-address>* / *<TAG>* } next-hop-self **no neighbor { *<ip-address>* / *<TAG>* } next-hop-self**	While sending route Next-Hop set Next-Hop as the switch's address; the **no neighbor { *<ip-address>* / *<TAG>* } next-hop-self** command cancels the setting.

2） Cancel default Next-Hop through route map

Command	Explanation
Route mapped configuration command	

set ip next-hop ***<ip-address>*** **no set ip next-hop**	Set the Next-Hop attribute of outbound route. The **no set ip next-hop** command cancels this setting.

7. Configure EGBP Multi-Hop

If the connections with outer neighbors are not direct, the following command can configure neighbor Multi-Hop.

Command	Explanation
BGP configuration mode	
neighbor {*<ip-address>*|*<TAG>*} ebgp-multihop [*<1-255>*] **no neighbor {*<ip-address>*|*<TAG>*} ebgp-multihop [*<1-255>*]**	Configure the allowance of EBGP connection with other networks that are not connected directly; the **no neighbor {*<ip-address>*|*<TAG>*} ebgp-multihop [*<1-255>*]** command cancels the setting.

8. Configure BGP session identifier

Command	Explanation
BGP configuration mode	
bgp router-id ***<ip-address>*** **no bgp router-id**	Configure the router-id value; the **no bgp router-id** command recovers the default value.

9. Configure the BGP Version

Command	Explanation
BGP configuration mode	
neighbor {*<ip-address>* | *<TAG>*} version <value> **no neighbor {*<ip-address>* | *<TAG>*} version**	Set the version used by BGP neighbors; the **no neighbor {*<ip-address>* | *<TAG>*} version** command recovers default setting. Presently only supporting version 4^{th}.

Ⅱ. Advanced BGP configuration tasks

1、 Use Route Maps to Modify Route

Command	Explanation
BGP configuration mode	
neighbor { *<ip-address>* | *<TAG>* } route-map *<map-name >* {in | out} **no neighbor { *<ip-address>* | *<TAG>* } route-map *<map-name >* {in | out}**	Apply a route map to incoming or outgoing routes; the **no neighbor { *<ip-address>* | *<TAG>* } route-map *<map-name >* {in | out}** command cancels the settings of routing maps.

2、 Configure Route Aggregation

Command	Explanation
BGP configuration mode	
aggregate-address <*ip-address/M*> [summary-only] [as-set] **no aggregate-address <*ip-address/M*> [summary-only] [as-set]**	Create an aggregate entry in the BGP routing table; the **no aggregate-address <*ip-address/M*> [summary-only] [as-set]** command cancels the aggregate entry.

3、 Configure BGP Community Filtering

Command	Explanation
BGP configuration mode	
neighbor {<*ip-address*> / <*TAG*>} send-community **no neighbor {<*ip-address*> / <*TAG*>} send-community**	Allow the routing updates with community attributes sending to BGP neighbors; the **no neighbor {<*ip-address*> / <*TAG*>} send-community** command enables the route without community attributes.

4、 Configure BGP Confederation

Command	Explanation
BGP configuration mode	
bgp confederation identifier <*as-id*> **no bgp confederation identifier <*as-id*>**	Configure a BGP AS confederation identifier; the **no bgp confederation identifier <*as-id*>** command deletes the BGP AS confederation identifier.
bgp confederation peers <*as-id*> [<*as-id*>..] **no bgp confederation peers <*as-id*> [<*as-id*>..]**	Configure the AS affiliated to the AS confederation; the **no bgp confederation peers <*as-id*> [<*as-id*>..]** command deletes the AS from the AS confederation.

5、 Configure a Route Reflector

（1） The following commands can be used to configure route reflector and its clients.

Command	Explanation
BGP configuration mode	

neighbor *<ip-address>* route-reflector-client **no neighbor *<ip-address>* route-reflector-client**	Configure the current switch as route reflector and specify a client; the **no neighbor *<ip-address>* route-reflector-client** command format deletes a client.

（2） **If there are more than one route reflectors in the cluster, the following commands can configure cluster-id**

Command	Explanation
BGP configuration mode	
bgp cluster-id *<cluster-id>* **no bgp cluster-id**	Configure cluster id; the **no bgp cluster-id** command cancels the cluster id configuration.

（3） **If the route reflector from clients to clients is needed, the following commands can be used.**

Command	Explanation
BGP configuration mode	
bgp client-to-client reflection **no bgp client-to-client reflection**	Configure the allowance of the route reflector from clients to clients; the **no bgp client-to-client reflection** command forbids this allowance.

6、 Configure Peer Groups

（1） **Create peer groups**

Command	Explanation
BGP configuration mode	
neighbor *<TAG>* peer-group **no neighbor *<TAG>* peer-group**	Create peer groups; the **no neighbor *<TAG>* peer-group** command deletes peer groups.

（2） **Add neighbors to peers groups**

Command	Explanation
BGP configuration mode	
neighbor *<ip-address>* peer-group *<TAG>* **no neighbor *<ip-address>* peer-group *<TAG>***	Make a neighbor a member of the peer group. The **no neighbor *<ip-address>* peer-group *<TAG>*** command cancels the specified member.

7、 Configure neighbors and peer Groups' parameters

Command	Explanation
BGP configuration mode	
**neighbor {*<ip-address>*	*<TAG>*} remote-as *<as-id>*** **no neighbor {*<ip-address>*	*<TAG>*} remote-as**	Specify a BGP neighbor; the **no neighbor {*<ip-address>*	*<TAG>*}**

<as-id>	**remote-as *<as-id>*** command deletes the neighbor.
neighbor { *<ip-address>* / *<TAG>* } description *<.LINE>* **no neighbor { *<ip-address>* / *<TAG>* } description**	Associate a description with a neighbor; the **no neighbor {*<ip-address>* / *<TAG>*} description** command deletes this description.
neighbor { *<ip-address>* / *<TAG>* } default-originate [route-map *<NAME>*] **no neighbor { *<ip-address>* / *<TAG>* } default-originate [route-map *<NAME>*]**	Permit to send the default route 0.0.0.0; the **no neighbor { *<ip-address>* / *<TAG>* } default-originate [route-map *<NAME>*]** command cancels sending default route.
neighbor { *<ip-address>* / *<TAG>* } send-community **no neighbor { *<ip-address>* / *<TAG>* } send-community**	Configure the community attributes sent to the neighbor.
neighbor { *<ip-address>* / *<TAG>* } timers *<keep alive> <holdtime>* **no neighbor { *<ip-address>* / *<TAG>* } timers**	Configure a particular neighbor's keep-alive and hold-time timer; the **no neighbor {*<ip-address>* / *<TAG>*} timers** command recovers the default value.
neighbor {*<ip-address>* / *<TAG>*} advertisement-interval *<seconds>* **no neighbor {*<ip-address>*	*<TAG>*} advertisement-interval**	Configure the min interval of sending BGP routing information; the **no neighbor {*<ip-address>*	*<TAG>*} advertisement-interval** command recovers the default value.
neighbor {*<ip-address>* / *<TAG>*} ebgp-multihop[*<1-255>*] **no neighbor {*<ip-address>* / *<TAG>*} ebgp-multihop**	Configure the allowance of EBGP connections with networks connected indirectly; the **no neighbor {*<ip-address>* / *<TAG>*} ebgp-multihop** command cancels this setting.
neighbor { *<ip-address>* / *<TAG>* } weight <weight> **no neighbor { *<ip-address>* / *<TAG>* } weight**	Configure BGP neighbor weights; the **no neighbor { *<ip-address>* / *<TAG>* } weight** command recovers the default weights.
**neighbor { *<ip-address>* / *<TAG>* } distribute-list { *<access-list-number>*	*<name>* } { in	out }** **no neighbor { *<ip-address>* / *<TAG>* } distribute-list { *<access-list-number>*	*<name>* } { *in*	*out* }**	Filter neighbor route update; the **no neighbor { *<ip-address>* / *<TAG>* } distribute-list { *<access-list-number>*	*<name>* } { *in*	*out* }** command cancels route filtering.
neighbor { *<ip-address>* / *<TAG>* } route-reflector-client **no neighbor { *<ip-address>* / *<TAG>* } route-reflector-client**	Configure the current switch as route reflector and specify a client; the **no neighbor { *<ip-address>* / *<TAG>* } route-reflector-client** command

	deletes a client.
neighbor { *<ip-address>* / *<TAG>* } next-hop-self **no neighbor { *<ip-address>* / *<TAG>* } next-hop-self**	When sending route, configure Next-Hop as its address; the **no neighbor { *<ip-address>* / *<TAG>* } next-hop-self** command cancels the setting.
neighbor { *<ip-address>* / *<TAG>* } version *<value>* **no neighbor { *<ip-address>* / *<TAG>* } version**	Specify the BGP version communicating with BGP neighbors; the **no neighbor { *<ip-address>* / *<TAG>* } version** command recovers default setting.
neighbor { *<ip-address>* / *<TAG>* } route-map *<map-name>* {in | out} **no neighbor { *<ip-address>* / *<TAG>* } route-map *<map-name>* {in | out}**	Apply a route map to incoming or outgoing routes; the **no neighbor { *<ip-address>* / *<TAG>* } route-map *<map-name>* {in | out}** command cancels the setting of route reflector.
neighbor { *<ip-address>* / *<TAG>* } soft-reconfiguration inbound **no neighbor { *<ip-address>* / *<TAG>* } soft-reconfiguration inbound**	Store the route information from neighbor or peers; the **no neighbor { *<ip-address>* / *<TAG>* } soft-reconfiguration inbound** command cancels the storage.
neighbor { *<ip-address>* / *<TAG>* } shutdown **no neighbor { *<ip-address>* / *<TAG>* } shutdown**	Shutdown BGP neighbor or peers; the **no neighbor { *<ip-address>* / *<TAG>* } shutdown** command activates the closed BGP neighbor or peers.

8、Adjust BGP Timers

（1） **Configure the BGP timer of all the neighbors**

Command	Explanation
BGP configuration mode	
timers bgp *<keep alive>* *<holdtime>* **no timers bgp**	Configure the BGP timers of all the neighbors; the no timer bgp command recovers the default value.

（2） **Configure the timer value of a particular neighbor**

Command	Explanation
BGP configuration mode	

neighbor { *<ip-address>* / *<TAG>* } timers *<keep alive>* *<holdtime>* **no neighbor { *<ip-address>* / *<TAG>* } timers**	Configure the keep alive and holdtime timer of a particular neighbor; the **no neighbor { *<ip-address>* / *<TAG>* } timers** command recovers the default value.

9、Adjust BGP announcement Interval

Command	Explanation
BGP configuration mode	
neighbor {*<ip-address>* / *<TAG>*} advertisement-interval <seconds> **no neighbor {*<ip-address>* / *<TAG>*} advertisement-interval**	Configure the minimum interval among BGP routes update information; the **no neighbor {*<ip-address>* / *<TAG>*} advertisement-interval** command recovers the default setting.

10、Configure the Local Preference Value

Command	Explanation
BGP configuration mode	
bgp default local-preference *<value>* **no bgp default local-preference**	Change default local preference; the **no bgp default local-preference** command recovers the default value.

11、Enable sending default route

Command	Explanation
BGP configuration mode	
neighbor { *<ip-address>* / *<TAG>* } default-originate **no neighbor { *<ip-address>* / *<TAG>* } default-originate**	Permit sending default route 0.0.0.0; the **no neighbor { *<ip-address>* / *<TAG>* } default-originate** command cancels sending default route.

12、Configure BGP's MED Value

（1） Configure MED value

Command	Explanation
Route map configuration command	
set metric *<metric-value>* **no set metric**	Configure metric value; the **no set metric** command recovers the default value.

（2） Apply route selection based on MED according to the path from different AS

Command	Explanation
BGP configuration mode	

bgp always-compare-med **no bgp always-compare-med**	Permit the MED comparison from different AS; the **no bgp always-compare-med** command forbids the comparison.

13、Configure BGP routing redistribution

Command	Explanation
BGP configuration mode	
redistribute { connected | static | rip | ospf} [metric *<metric>*] [route-map *<NAME>*] **no redistribute { connected | static | rip | ospf}**	Redistribute IGP routes to BGP and may specify the redistributed metric and route reflector; the **no redistribute { connected | static | rip | ospf}** command cancels the redistribution.

14、Configure Route Dampening

Command	Explanation
BGP configuration mode	
bgp dampening [*<1-45>*] [*<1-20000>* *<1-20000>* *<1-255>*] [*<1-45>*] **no bgp dampening**	Enable BGP route dampening and apply the specified parameters; the **no bgp dampening** command stops route dampening

15、Configure BGP capability Negotiation

Command	Explanation
BGP configuration mode	
neighbor {*<ip-address>*|*<TAG>*} capability {dynamic | route-refresh} **no neighbor {*<ip-address>*|*<TAG>*} capability {dynamic | route-refresh}** **neighbor {*<ip-address>*|*<TAG>*} capability orf prefix-list {*<both>*|*<send>*|*<receive>*}** **no neighbor {*<ip-address>*|*<TAG>*} capability orf prefix-list {*<both>*|*<send>*|*<receive>*}** **neighbor {*<ip-address>*|*<TAG>*} dont-capability-negotiate** **no neighbor {*<ip-address>*|*<TAG>*} dont-capability-negotiate** **neighbor {*<ip-address>*|*<TAG>*} override-capability** **no neighbor {*<ip-address>*|*<TAG>*} override-capability** **neighbor {*<ip-address>*|*<TAG>*} strict-capability-match** **no neighbor {*<ip-address>*|*<TAG>*} strict-capability-match**	BGP provides capability negotiation regulation and carry out this capability match while establishing connection. The currently supported capabilities include route update, dynamic capability, outgoing route filtering capability and the address family's capability of supporting the negotiation. Use these command to enable these capabilities, its format "no" close these capabilities .It can also be configured by commands to not do capability negotiation, do strict capability negotiation or not care about the negotiation results.

16。 Configure Routing Server

Command	Explanation
BGP configuration mode	
neighbor {*<ip-address>*|*<TAG>*} route-server-client **no neighbor {*<ip-address>*|*<TAG>*} route-server-client**	Route server may configure BGP neighbors under EBGP environment to reduce the number of peers that every client has configured; format "no" of the command configures this router as route server and specify the clients it serves, the **no neighbor {*<ip-address>*|*<TAG>*} route-server-client** command can delete clients.

17。 Configure Path-selected rules

Command	Explanation
BGP configuration mode	
bgp always-compare-med **no bgp always-compare-med** **bgp bestpath as-path ignore** **no bgp bestpath as-path ignore** **bgp bestpath compare-confed-aspath** **no bgp bestpath compare-confed-aspath** **bgp bestpath compare-routerid** **no bgp bestpath compare-routerid** **bgp bestpath med {[confed] [missing-is-worst]}** **no bgp bestpath med {[confed] [missing-is-worst]}**	BGP may change some path-select rules by configuration to change the best selection and compare MED under EBGP environment through these command, ignore the AS-PATH length, compare the confederation as-path length, compare the route identifier and compare the confederation MED etc. Its format "no" recovers the default route path-selected rules.

18. Configure redistribution of OSPF routing to BGP

(1) Enable redistribution of OSPF routing to BGP

Command	Explanation
Router BGP Configuration Mode	
redistribute ospf [*<process-id>*] [route-map*<word>*] **no redistribute ospf [<process-id>]**	To enable or disable the redistribution of OSPF routing to BGP.

(2)Display and debug the information about configuration of redistribution of OSPF routing to BGP

Command	Explanation
Admin Mode and Configuration Mode	
show ip bgp redistribute	To enable or disable the redistribution of OSPF routing to BGP.
Admin Mode	

debug bgp redistribute message send **no debug bgp redistribute message send** **debug bgp redistribute route receive** **no debug bgp redistribute route receive**	To enable or disable debugging messages sent by BGP for redistributing OSPF routing. To enable or disable debugging messages received from NSM for redistributing OSPF routing.

4.5.3 Configuration Examples of BGP

4.5.3.1 Examples 1: configure BGP neighbor

SwitchB, SwitchC and SwitchD are in AS200, SwitchA is in AS100. SwitchA and SwitchB share the same network segment. SwitchB and SwitchD are not connected physically.

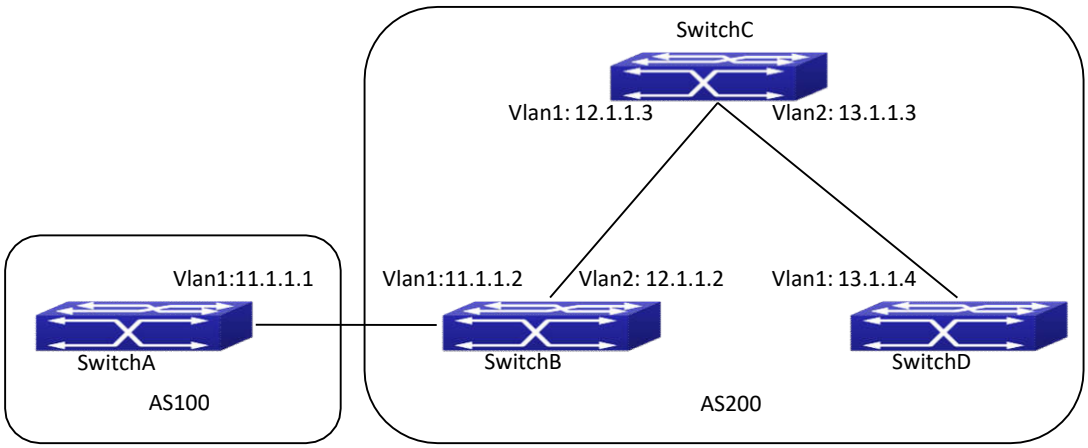

Figure 4-9 BGP Network Topological Map

The configurations of SwitchA are as following:
SwitchA(config)#router bgp 100
SwitchA(config-router-bgp)#neighbor 11.1.1.2 remote-as 200
SwitchA(config-router-bgp)#exit

The configurations of SwitchB are as following:
SwitchB(config)#router bgp 200
SwitchB(config-router-bgp)#network 11.0.0.0
SwitchB(config-router-bgp)#network 12.0.0.0
SwitchB(config-router-bgp)#network 13.0.0.0
SwitchB(config-router-bgp)#neighbor 11.1.1.1 remote-as 100
SwitchB(config-router-bgp)#neighbor 12.1.1.3 remote-as 200
SwitchB(config-router-bgp)#neighbor 13.1.1.4 remote-as 200
SwitchB(config-router-bgp)#exit

The configurations of SwitchC are as following:

SwitchC(config)#router bgp 200
SwitchC(config-router-bgp)#network 12.0.0.0
SwitchC(config-router-bgp)#network 13.0.0.0
SwitchC(config-router-bgp)#neighbor 12.1.1.2 remote-as 200
SwitchC(config-router-bgp)#neighbor 13.1.1.4 remote-as 200
SwitchC(config-router-bgp)#exit

The configurations of SwitchD are as following:
SwitchD(config)#router bgp 200
SwitchD(config-router-bgp)#network 13.0.0.0
SwitchD(config-router-bgp)#neighbor 12.1.1.2 remote-as 200
SwitchD(config-router-bgp)#neighbor 13.1.1.3 remote-as 200
SwitchD(config-router-bgp)#exit

Presently, the connection between SwitchB and SwitchA is EBGP, and other connections with SwitchC and SwitchD are IBGP. SwitchB and SwitchD may have BGP connection without physical connection. But there is a precondition that these two switches must have reachable route to each other. This route can be attained through static route or IGP.

4.5.3.2 Examples 2: configure BGP aggregation

In this sample, configure route aggregation. Firstly, enable command redistribute to redistribute static route to BGP route table:
SwitchB(config)#ip route 193.0.0.0/24 11.1.1.1
SwitchB(config)#router bgp 100
SwitchB(config-router-bgp)#redistribute static
When there is at least one route affiliated to the specified range, the following configuration will create an aggregation route in the BGP route table. The aggregation route will be regarded as the AS from itself. More detailed route information about 193.0.0.0 will be announced.
SwitchB(config#router bgp 100
SwitchB(config-router-bgp)#aggregate 193.0.0.0/16
At the same time, the aggregation command above can be modified as following, then this switch only announce aggregation route 193.0.0.0 and forbid to announce more specified route to all the neighbors.
SwitchB(config-router-bgp)#aggregate 193.0.0.0/16 summary-only

4.5.3.3 Examples 3: configure BGP community attributes

In the following sample, “route map set-community” is used for the outgoing update to neighbor 16.1.1.6. By accessing to route in table 1 to configure special community value to “1111”, other can be announced normally.
Switch(config)#router bgp 100
Switch(config-router-bgp)#neighbor 16.1.1.6 remote-as 200
Switch(config-router-bgp)#neighbor 16.1.1.6 route-map set-community out
Switch(config-router-bgp)#exit

```
Switch(config)#route-map set-community permit 10
Switch(config-route-map)#match address 1
Switch(config-route-map)#set community 1111
Switch(config-route-map)#exit
Switch(config)#route-map set-community permit 20
Switch(config-route-map)#match address 2
Switch(config-route-map)#exit
Switch(config)#access-list 1 permit 11.1.0.0 0.0.255.255
Switch(config)#access-list 2 permit 0.0.0.0 255.255.255.255
Switch(config)#exit
Switch#clear ip bgp 16.1.1.6 soft out
```

In the following sample, configure the MED local preference of the routes from neighbor 16.1.1.6 selectively according to the route community value. All the routes that match the community list will set MED as 2000, community list com1 permits the route with community value “100 200 300”or”900 901” to pass. This route may have other community attributes. All the routes that pass community list com2 will set the local preference as 500. But the route that can’t pass both com1 and com2 will be rejected.

```
Switch(config)#router bgp 100
Switch(config-router-bgp)#neighbor 16.1.1.6 remote-as 200
Switch(config-router-bgp)#neighbor 16.1.1.6 route-map match-community in
Switch(config-router-bgp)#exit
Switch(config)#route-map match-community permit 10
Switch(config-route-map)#match community com1
Switch(config-route-map)#set metric 2000
Switch(config-route-map)#exit
Switch(config)#route-map match-community permit 20
Switch(config-route-map)#match community com2
Switch(config-route-map)#set local-preference 500
Switch(config-route-map)#exit
Switch(config)#ip community-list com1 permit 100 200 300
Switch(config)#ip community-list com1 permit 900 901
Switch(config)#ip community-list com2 permit 88
Switch(config)#ip community-list com2 permit 90
Switch(config)#exit
Switch#clear ip bgp 16.1.1.6 soft out
```

4.5.3.4 Examples 4: configure BGP confederation

The following is the configuration of an AS. As the picture illustrated, SwitchB and SwitchC establish IBGP connection. SwitchD is affiliated to AS 20.SwitchB and SwitchC establish EBGP of inner AS confederation. AS10 and AS20 form AS confederation with the AS number AS200; SwitchA belongs to AS100, SwitchB may create EBGP connection by AS200.

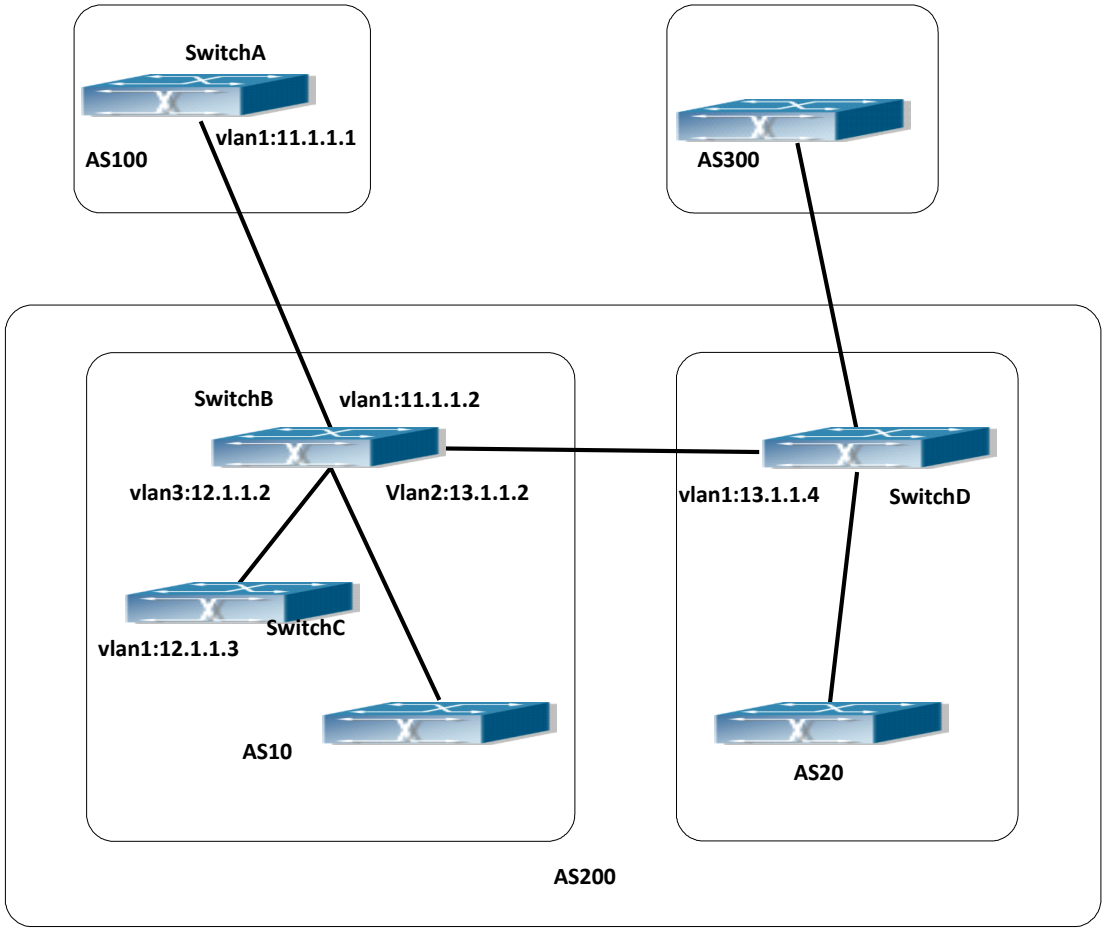

Figure 4-10 Confederation configuring topology

The configurations are as following:

SwitchA:

SwitchA(config)#router bgp 100

SwitchA(config-router-bgp)#neighbor 11.1.1.2 remote-as 200

SwitchB:

SwitchB(config)#router bgp 10

SwitchB(config-router-bgp)#bgp confederation identifier 200

SwitchB(config-router-bgp)#bgp confederation peers 20

SwitchB(config-router-bgp)#neighbor 12.1.1.3 remote-as 10

SwitchB(config-router-bgp)#neighbor 13.1.1.4 remote-as 20

SwitchB(config-router-bgp)#neighbor 11.1.1.1 remote-as 100

SwitchC:

SwitchC(config)#router bgp 10

SwitchC(config-router-bgp)#bgp confederation identifier 200

SwitchC(config-router-bgp)#bgp confederation peers 20

SwitchC(config-router-bgp)#neighbor 12.1.1.2 remote-as 10

SwitchD:

SwitchD(config)#router bgp 20
SwitchD(config-router-bgp)#bgp confederation identifier 200
SwitchD(config-router-bgp)#bgp confederation peers 10
SwitchD(config-router-bgp)#neighbor 13.1.1.2 remote-as 10

4.5.3.5 Examples 5: configure BGP route reflector

The following is the configuration of a route reflector. As the picture illustrated, SwitchA, SwitchB, SwitchC, SwitchD, SWE, SWF and SWG establish IBGP connection which is affiliated to AS100. SwitchC creates EBGP connection with AS200. SwitchA creates EBGP connection with AS300. SwitchC, SwitchD and SWG make route reflectors.

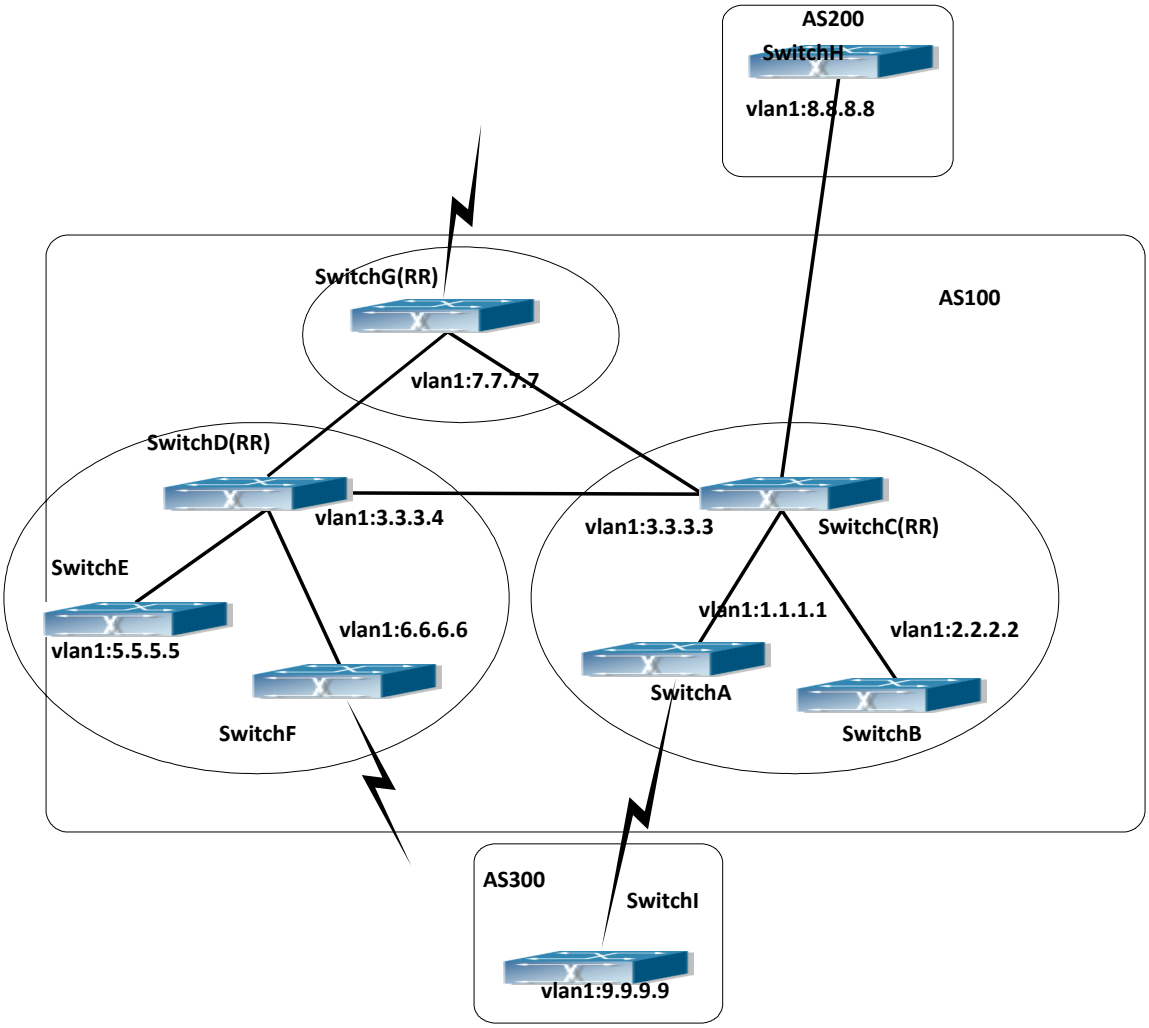

Figure 4-11 the Topological Map of Route Reflector

The configurations are as following:
The configurations of SwitchC:
SwitchC(config)#router bgp 100
SwitchC(config-router-bgp)#neighbor 1.1.1.1 remote-as 100

SwitchC(config-router-bgp)#neighbor 1.1.1.1 route-reflector-client
SwitchC(config-router-bgp)#neighbor 2.2.2.2 remote-as 100
SwitchC(config-router-bgp)#neighbor 2.2.2.2 route-reflector-client
SwitchC(config-router-bgp)#neighbor 7.7.7.7 remote-as 100
SwitchC(config-router-bgp)#neighbor 3.3.3.4 remote-as 100
SwitchC(config-router-bgp)#neighbor 8.8.8.8 remote-as 200

The configurations of SwitchD:
SwitchD(config)#router bgp 100
SwitchD(config-router-bgp)#neighbor 5.5.5.5 remote-as 100
SwitchD(config-router-bgp)#neighbor 5.5.5.5 route-reflector-client
SwitchD(config-router-bgp)#neighbor 6.6.6.6 remote-as 100
SwitchD(config-router-bgp)#neighbor 6.6.6.6 route-reflector-client
SwitchD(config-router-bgp)#neighbor 3.3.3.3 remote-as 100
SwitchD(config-router-bgp)#neighbor 7.7.7.7 remote-as 100

The configurations of SwitchA:
SwitchA(config)#router bgp 100
SwitchA(config-router-bgp)#neighbor 1.1.1.2 remote-as 100
SwitchA(config-router-bgp)#neighbor 9.9.9.9 remote-as 300

The SwitchA at this time needn't to create IBGP connection with all the switches in the AS100 and could receive BGP route from other switches in the AS.

4.5.3.6 Examples 6: configure MED of BGP

The following is the configuration of a MED. As illustrated, SwitchA is affiliated to AS100, SwitchB is affiliated to AS400, SwitchC and SwitchD belong to AS300.

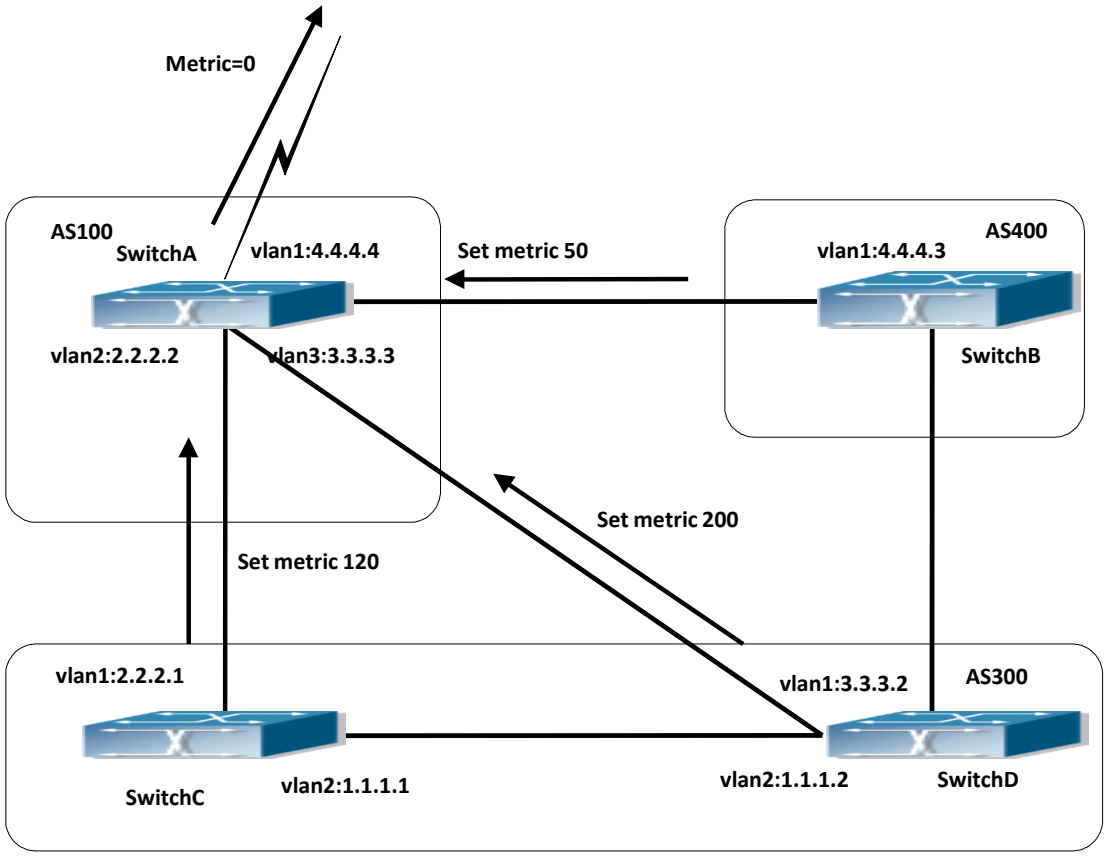

Figure 4-12 MED Configuring Topological Map

The configurations of SwitchA:

SwitchA(config)#router bgp 100

SwitchA(config-router-bgp)#neighbor 2.2.2.1 remote-as 300

SwitchA(config-router-bgp)#neighbor 3.3.3.2 remote-as 300

SwitchA(config-router-bgp)#neighbor 4.4.4.3 remote-as 400

The configurations of SwitchC:

SwitchC(config)#router bgp 300

SwitchC (config-router-bgp)#neighbor 2.2.2.2 remote-as 100

SwitchC (config-router-bgp)#neighbor 2.2.2.2 route-map set-metric out

SwitchC (config-router-bgp)#neighbor 1.1.1.2 remote-as 300

SwitchC (config-router-bgp)#exit

SwitchC (config)#route-map set-metric permit 10

SwitchC (Config-Router-RouteMap)#set metric 120

The configurations of SwitchD

SwitchD (config)#router bgp 300

SwitchD (config-router-bgp)#neighbor 3.3.3.3 remote-as 100

SwitchD (config-router-bgp)#neighbor 3.3.3.3 route-map set-metric out

SwitchD (config-router-bgp)#neighbor 1.1.1.1 remote-as 300

SwitchD (config-router-bgp)#exit

SwitchD (config)#route-map set-metric permit 10
SwitchD (Config-Router-RouteMap)#set metric 200

The configurations of SwitchB
SwitchB (config)#router bgp 400
SwitchB (config-router-bgp)#neighbor 4.4.4.4 remote-as 100
SwitchB (config-router-bgp)#neighbor 4.4.4.4 route-map set-metric out
SwitchB (config-router-bgp)#exit
SwitchB (config)#route-map set-metric permit 10
SwitchB (Config-Router-RouteMap)#set metric 50

After the configuration above, SwitchB, SwitchC and SwitchD are assumed to send a route 12.0.0.0 to SwitchA. According to the comparison of BGP route strategy; there is an assumption that the routes sent by the three switches above have the same attribute value before the comparison of metric attribute. At this time, the route with lower value is the better route. But the comparison of metric attribute will only be done with the routes from the same AS. For SwitchA, the routes passed SwitchC are preferable to the one passed SwitchD. Because SwitchC and SwitchB are not located in the same AS, the SwitchA will not do metric comparison between the two switches. If the metric comparison between different AS is needed, the command" bgp always-compare-med" will be used. If this command is configured, the routes passed SwitchB are the best to SwitchA. At this time, the following command may be added on SwitchA:
SwitchA (config-router-bgp)#bgp always-compare-med

4.5.3.7 Examples 7: example of BGP VPN

For the configuration of MPLS VPN, BGP is part of the core routing system and it is also an important utility to support ILM and FTN entries on the edge devices. For switch, the BGP protocol together with the LDP protocol, constructs the foundation of the MPLS VPN application. The LDP protocol works at the WLAN side and for the routers which are not on the edge of the network, the BGP protocol does not function.

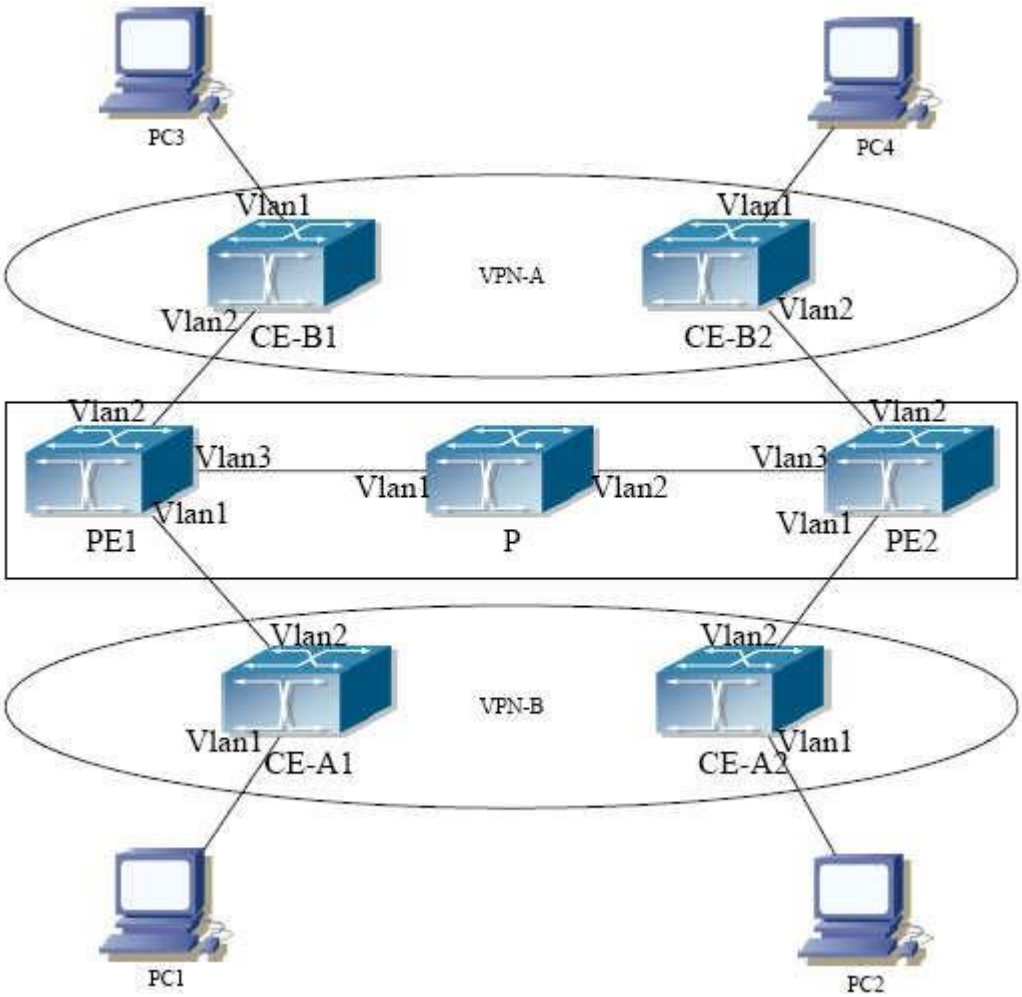

Figure 4-13 Example of MPLS VPN

As the figure shows, for a typical MPLS VPN application, the public network region consists of PE1, P and PE2, which MPLS is applied for packet transmission. VPN-A consists of CE-A1 and CE-A2, and VPN-B consists of CE-B1 and CE-B2. These two VPNs are isolated from each other. PE1 and PE2 are edge routers which are provided by the operators. CE-A1, CE-A2, CE-B1 and CE-B2 are the access switches on the user side. PC1-PC4 indicate the network users. BGP runs at both the public and private network region. For the public network region, VPN routing should be supported and the LOOPBACK interface should be used for connections.

The sample configurations are listed as below.

Configurations on CE-A1：

```
CE-A1#config
CE-A1(config)#interface vlan 2
CE-A1(config-if-Vlan2)#ip address 192.168.101.2 255.255.255.0
CE-A1(config-if-Vlan2)#exit
CE-A1(config)#interface vlan 1
CE-A1(config-if-Vlan2)#ip address 10.1.1.1 255.255.255.0
CE-A1(config-if-Vlan2)#exit
CE-A1(config)#router bgp 60101
```

```
CE-A1(config-router)#neighbor 192.168.101.1 remote-as 100
CE-A1(config-router)#exit
```

Configurations on CE-A2：.

```
CE-A2#config
CE-A2(config)#interface vlan 2
CE-A2(config-if-Vlan2)#ip address 192.168.102.2 255.255.255.0
CE-A2(config-if-Vlan2)#exit
CE-A2(config)#interface vlan 1
CE-A2(config-if-Vlan2)#ip address 10.1.2.1 255.255.255.0
CE-A2(config-if-Vlan2)#exit
CE-A2(config)#router bgp 60102
CE-A2(config-router)#neighbor 192.168.102.1 remote-as 100
CE-A2(config-router)#exit
```

Configurations on CE-B1：.

```
CE-B1#config
CE-B1(config)#interface vlan 2
CE-B1(config-if-Vlan2)#ip address 192.168.201.2 255.255.255.0
CE-B1(config-if-Vlan2)#exit
CE-B1(config)#interface vlan 1
CE-B1(config-if-Vlan2)#ip address 20.1.1.1 255.255.255.0
CE-B1(config-if-Vlan2)#exit
CE-B1(config)#router bgp 60201
CE-B1(config-router)#neighbor 192.168.201.1 remote-as 100
CE-B1(config-router)#exit
```

Configurations on CE-BE2：.

```
CE-B2#config
CE-B2(config)#interface vlan 2
CE-B2(config-if-Vlan2)#ip address 192.168.202.2 255.255.255.0
CE-B2(config-if-Vlan2)#exit
CE-B2(config)#interface vlan 1
CE-B2(config-if-Vlan2)#ip address 20.1.2.1 255.255.255.0
CE-B2(config-if-Vlan2)#exit
CE-B2(config)#router bgp 60202
CE-B2(config-router)#neighbor 192.168.202.1 remote-as 100
CE-B2(config-router)#exit
```

Configurations on PE1：

```
PE1#config
PE1(config)#ip vrf VRF-A
PE1(config-vrf)#rd 100:10
PE1(config-vrf)#route-target both 100:10
```

```
PE1(config-vrf)#exit
PE1(config)#ip vrf VRF-B
PE1(config-vrf)#rd 100:20
PE1(config-vrf)#route-target both 100:20
PE1(config-vrf)#exit
PE1(config)#interface vlan 1
PE1(config-if-Vlan1)#ip vrf forwarding VRF-A
PE1(config-if-Vlan1)#ip address 192.168.101.1 255.255.255.0
PE1(config-if-Vlan1)#exit
PE1(config)#interface vlan 2
PE1(config-if-Vlan2)#ip vrf forwarding VRF-B
PE1(config-if-Vlan2)#ip address 192.168.201.1 255.255.255.0
PE1(config-if-Vlan2)#exit
PE1(config)#interface vlan 3
PE1(config-if-Vlan3)#ip address 202.200.1.2 255.255.255.0
PE1(config-if-Vlan3)#label-switching
PE1(config-if-Vlan3)#exit
PE1(config)#interface loopback 1
PE1(Config-if-Loopback1)# ip address 200.200.1.1 255.255.255.255
PE1(config-if-Vlan3)#exit
PE1(config)#router bgp 100
PE1(config-router)#neighbor 200.200.1.2 remote-as 100
PE1(config-router)#neighbor 200.200.1.2 update-source 200.200.1.1
PE1(config-router)#address-family vpnv4 unicast
PE1(config-router-af)#neighbor 200.200.1.2 activate
PE1(config-router-af)#exit-address-family
PE1(config-router)#address-family ipv4 vrf VRF-A
PE1(config-router-af)# neighbor 192.168.101.2 remote-as 60101
PE1(config-router-af)#exit-address-family
PE1(config-router)#address-family ipv4 vrf VRF-B
PE1(config-router-af)# neighbor 192.168.201.2 remote-as 60201
PE1(config-router-af)#exit-address-family
```

Configurations on PE2：

```
PE2#config
PE2(config)#ip vrf VRF-A
PE2(config-vrf)#rd 100:10
PE2(config-vrf)#route-target both 100:10
PE2(config-vrf)#exit
PE2(config)#ip vrf VRF-B
PE2(config-vrf)#rd 100:20
PE2(config-vrf)#route-target both 100:20
PE2(config-vrf)#exit
PE2(config)#interface vlan 1
```

```
PE2(config-if-Vlan1)#ip vrf forwarding VRF-A
PE2(config-if-Vlan1)#ip address 192.168.102.1 255.255.255.0
PE2(config-if-Vlan1)#exit
PE2(config)#interface vlan 2
PE2(config-if-Vlan2)#ip vrf forwarding VRF-B
PE2(config-if-Vlan2)#ip address 192.168.202.1 255.255.255.0
PE2(config-if-Vlan2)#exit
PE2(config)#interface vlan 3
PE2(config-if-Vlan3)#ip address 202.200.2.2 255.255.255.0
PE2(config-if-Vlan3)#label-switching
PE2(config-if-Vlan3)#exit
PE2(config)#interface loopback 1
PE2(Config-if-Loopback1)# ip address 200.200.1.2 255.255.255.255
PE2(config-if-Vlan3)#exit
PE2(config)#router bgp 100
PE2(config-router)#neighbor 200.200.1.1 remote-as 100
PE2(config-router)#address-family vpnv4 unicast
PE2(config-router-af)#neighbor 200.200.1.1 activate
PE2(config-router-af)#exit-address-family
PE2(config-router)#address-family ipv4 vrf VRF-A
PE2(config-router-af)# neighbor 192.168.102.2 remote-as 60102
PE2(config-router-af)#exit-address-family
PE2(config-router)#address-family ipv4 vrf VRF-B
PE2(config-router-af)# neighbor 192.168.202.2 remote-as 60202
PE2(config-router-af)#exit-address-family
```

The sample configurations which are listed above is the most typical one. To enable communication between VRF, the route-target should be modified. And if the BGP AS number duplicates for the ends, the "**neighbor *<ip-addr>* as-override"** command should be configured to avoid the duplication of AS numbers.

Also, only BGP related configuration are listed above, to run LDP on the public network region, please refer to the LDP configuration sample.

4.5.4 BGP Troubleshooting

In the process of configuring and implementing BGP protocol, physical connection, configuration false probably leads to BGP protocol doesn't work. Therefore, the customers should give their attention to points as follow:

- ☞ First of all, to ensure correct physical connection;
- ☞ Secondly, to ensure interface and link protocol are UP (execute **show interface** instruction);
- ☞ And startup BGP protocol (use **router bgp** command), configure affiliated IBGP and EBGP neighbors (use **neighbor remote-as** command).

☞ Notice BGP protocol itself can't detect route, needs to import other routes to create BGP route. Only it enables these routes to announce IBGP and EBGP neighbors by importing routes. Direct-link routes, static route, and IGP route (RIP and OSPF) are included in these imported routes. **network** and **redistribute (BGP)** command are the ways of imported routes.

☞ For BGP, pay attention to the difference between the behaviors of IBGP and EBGP.

☞ After configuration finishes, the command of **show ip bgp summary** can be used to observe neighbor's connections, so that all of the neighbors keep BGP connection situation. And use **show ip bgp** command to observe BGP routing table.

☞ If BGP routing problem still can't be solved by debugging, please use debug instructions like **debug ip bgp** packet/events etc, and copy DEBUG information in 3 minutes, then send them to ourTechnology Service Center.

4.6 IPv4 Black Hole Routing

4.6.1 Introduction to Black Hole Routing

Black Hole Routing is a special kind of static routing which drops all the datagrams that match the routing rule.

4.6.2 IPv4 Black Hole Routing Configuration Task

1. Configure IPv4 Black Hole Routing

1. Configure IPv4 Black Hole Routing

Command	Explaination
Global Configuration Mode	
ip route {*<ip-prefix> <mask>*|*<ip-prefix>/<prefix-length>*} null0 [*<distance>*] **no ip route {*<ip-prefix> <mask>*|*<ip-prefix>/<prefix-length>*} null0**	To configure the static Black Hole Routing. The no form of this command will remove the specified Black Hole Routing configuration.

4.6.3 Black Hole Routing Configuration Exmaples

Example 1: IPv4 Black Hole Routing function.

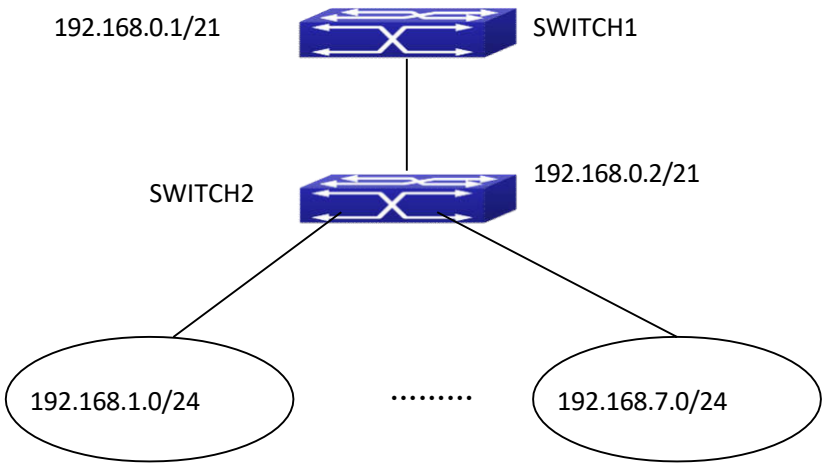

Figure 4-14 IPv4 Black Hole Routing Configuration Example

As it is shown in the figure, in Switch 2, eight in all interfaces are configured as Layer 3 VLAN interfaces for access interfaces. The network addresses are 192.168.1.0/24 ~ 192.268.7.0/24. A default routing is configured on Switch 2 to connect to Switch 1. And a backward default routing is configured on Switch 1 to Switch 2, whose network address is 192.168.0.0/21. Commonly, this configuration will work well. However, if one of the Layer 3 interfaces in Switch 2 goes down, for example, the interface belonged to 192.168.1.0/24. When datagrams arrives at VLAN1 in Switch 2, there will be no routing rules for these datagrams. The switch then will forward these datagrams according to the default routing, back to Switch 1. When Switch 1 receives these datagrams, it will forward them back to Switch 2. Thus, loopback exists. To solve this problem, Black Hole Routing can be introduced on Switch 2.

ip route 192.168.0.0/21 null0 50

Then Switch 2 will drop the datagrams from interface VLAN1 that match the Black Hole Routing rule. And loopback routing is prevented.

Configuration steps are listed as below:

Switch#config

Switch(config)#ip route 192.168.0.0/21 null0 50

4.6.4 Black Hole Routing Troubleshooting

When configuring the Black Hole Routing function, the configuration may not work due to some reasons such as incorrect network address mask, and incorrect management distance. Attention should be paid to the following items:

- ☞ IPv6 should be enabled before IPv6 Black Hole Routing can work.
- ☞ It is suggested that the length of the network address mask should be longer than that of normal routing configuration, in order to prevent the Black Hole Routing from intervening other routing configuration.
- ☞ When the network address mask of Black Hole Routing configuration is the same with some

other configuration, it is suggested that the distance of Black Hole Routing is set lower.

For problems that cannot be fixed through above methods, please issue the command show ip route distance and show ip route fib, and show l3. And copy and paste the output of the commands, and send to the technical service center of our company.

4.7 ECMP

4.7.1 Introduction to ECMP

ECMP (Equal-cost Multi-path Routing) works in the network environment where there are many different links to arrive at the same destination address. If using the traditional routing technique, only a link can be used to send the data packets to the destination address, other links at the backup state or the invalidation state, and it needs some times to process the mutual switchover under the static routing environment. However, ECMP protocol can use multi-links under such network environment, it not only implements the load balance, increases the transport bandwidth, but also can completely backup the data transport of the invalidation links without delay and packet loss.

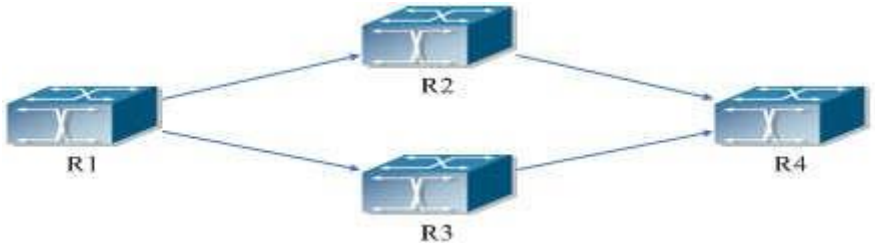

Figure 4-15 the application environment of ECMP

As it is shown in the figure, there are two paths can be selected from R1 to R4, they are R1-R2-R4 and R1-R3-R4. If the route type and the cost are same, then it can forms two routes from R1 to R4, but the next hop is different. If two routes are selected as the best, then they form the equal-cost route.

4.7.2 ECMP Configuration Task List

1. Configure the max number of equal-cost route
2. Configure load-balance mode for port-group

1. Configure the max number of equal-cost route

Command	Explanation
Global mode	
maximum-paths <1-8> **no maximum-paths**	Configure the max number of equal-cost route.

2. Configure load-balance mode for port-group

Command	Explanation
Global Mode	
load-balance {dst-src-mac | dst-src-ip | dst-src-mac-ip }	Set load-balance for switch, it takes effect for port-group and ECMP function at the same time.

4.7.3 ECMP Typical Example

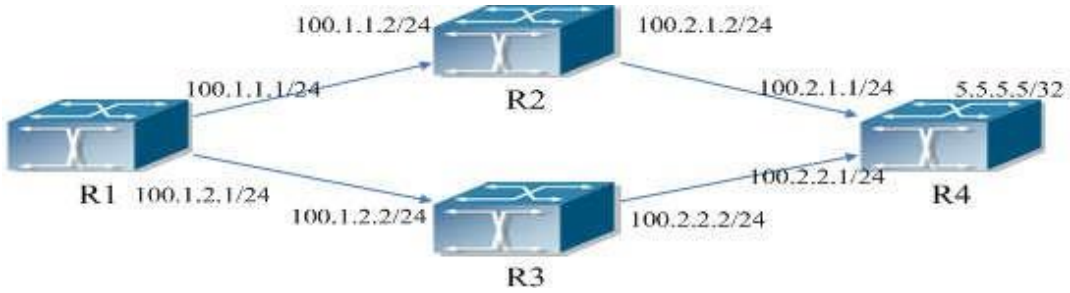

Figure 4-16 the application environment of ECMP

As it is shown in the figure, the R1 connect to R2 and R3 with the interface address 100.1.1.1/24 and 100.1.2.1/24. The R2 and R3 connect to R1 with the interface address 100.1.1.2/24 and 100.1.2.2/24. The R4 connect to R2 and R3 with interface address 100.2.1.1/24 and 100.2.2.1/24. The R2 and R3 connect to R4 with the interface address 100.2.1.2/24, 100.2.2.2/24. The loopback address of R4 is 5.5.5.5/32.

4.7.3.1 Static Route Implements ECMP

R1(config)#ip route 5.5.5.5/32 100.1.1.2
R1(config)#ip route 5.5.5.5/32 100.1.2.2
On R1, show ip route, the following is displayed:
R1(config)#show ip route
Codes: K - kernel, C - connected, S - static, R - RIP, B - BGP
O - OSPF, IA - OSPF inter area

```
        N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
        E1 - OSPF external type 1, E2 - OSPF external type 2
        i - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2, ia - IS-IS inter area
        * - candidate default
C        1.1.1.1/32 is directly connected, Loopback1     tag:0
S        5.5.5.5/32 [1/0] via 100.1.1.2, Vlan100     tag:0
                          [1/0] via 100.1.2.2, Vlan200     tag:0
C        100.1.1.0/24 is directly connected, Vlan100     tag:0
C        100.1.2.0/24 is directly connected, Vlan200 tag:0
C        127.0.0.0/8 is directly connected, Loopback     tag:0
                                    Total routes are : 6 item(s)
```

4.7.3.2 OSPF Implements ECMP

R1 configuration:

```
R1(config)#interface Vlan100
R1(Config-if-Vlan100)# ip address 100.1.1.1 255.255.255.0
R1(config)#interface Vlan200
R1(Config-if-Vlan200)# ip address 100.1.2.1 255.255.255.0
R1(config)#interface loopback 1
R1(Config-if-loopback1)# ip address 1.1.1.1 255.255.255.255
R1(config)#router ospf 1
R1(config-router)# ospf router-id 1.1.1.1
R1(config-router)# network 100.1.1.0/24 area 0
R1(config-router)# network 100.1.2.0/24 area 0
```

R2 configuration:

```
R2(config)#interface Vlan100
R2(Config-if-Vlan100)# ip address 100.1.1.2 255.255.255.0
R2(config)#interface Vlan200
R2(Config-if-Vlan200)# ip address 100.2.1.2 255.255.255.0
R2(config)#interface loopback 1
R2(Config-if-loopback1)# ip address 2.2.2.2 255.255.255.255
R2(config)#router ospf 1
R2(config-router)# ospf router-id 2.2.2.2
R2(config-router)# network 100.1.1.0/24 area 0
R2(config-router)# network 100.2.1.0/24 area 0
```

R3 configuration:

```
R3(config)#interface Vlan100
R3(Config-if-Vlan100)# ip address 100.1.2.2 255.255.255.0
R3(config)#interface Vlan200
R3(Config-if-Vlan200)# ip address 100.2.2.2 255.255.255.0
R3(config)#interface loopback 1
```

```
R3(Config-if-loopback1)# ip address 3.3.3.3 255.255.255.255
R3(config)#router ospf 1
R3(config-router)# ospf router-id 3.3.3.3
R3(config-router)# network 100.1.2.0/24 area 0
R3(config-router)# network 100.2.2.0/24 area 0

R4 configuration:
R4(config)#interface Vlan100
R4(Config-if-Vlan100)# ip address 100.2.1.1 255.255.255.0
R4(config)#interface Vlan200
R4(Config-if-Vlan200)# ip address 100.2.2.1 255.255.255.0
R4(config)#interface loopback 1
R4(Config-if-loopback1)# ip address 5.5.5.5 255.255.255.255
R4(config)#router ospf 1
R4(config-router)# ospf router-id 4.4.4.4
R4(config-router)# network 100.2.1.0/24 area 0
R4(config-router)# network 100.2.2.0/24 area 0

On R1, show ip route, the following is displayed:
R1(config)#show ip route
Codes: K - kernel, C - connected, S - static, R - RIP, B - BGP
       O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       i - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2, ia - IS-IS inter area
       * - candidate default
C        1.1.1.1/32 is directly connected, Loopback1    tag:0
O        5.5.5.5/32 [110/3] via 100.1.1.2, Vlan100, 00:00:05    tag:0
                    [110/3] via 100.1.2.2, Vlan200, 00:00:05    tag:0
C        100.1.1.0/24 is directly connected, Vlan100    tag:0
C        100.1.2.0/24 is directly connected, Vlan200    tag:0
O        100.2.1.0/24 [110/2] via 100.1.1.2, Vlan100, 00:02:25    tag:0
O        100.2.2.0/24 [110/2] via 100.1.2.2, Vlan200, 00:02:25    tag:0
C        127.0.0.0/8 is directly connected, Loopback    tag:0
                              Total routes are : 8 item(s)
```

4.7.4 ECMP Troubleshooting

When configuring ECMP, ECMP may not run normally for the reasons of physical connection and false configuration, so users should note the following essential.

☞ When using ECMP, load-balance mode should be set as dst-src-ip or dst-src-mac-ip, after that, load-balance is correct for packets.

4.8 BFD

4.8.1 Introduction to BFD

BFD (Bidirectional Forwarding Detection) provides a detection mechanism to quickly detect and monitor the connectivity of links in networks. To improve network performance, between protocol neighbors must quickly detect communication failures to restore communication through backup paths as soon as possible.

BFD provides a general-purpose, standard, medium-independent and protocol-independent fast failure detection mechanism. It can uniformly and quickly detect the failures of the bidirectional forwarding paths between two network devices for superstratum protocols, such as routing protocols and Multiprotocol Label Switching (MPLS). BFD establishes session between two network devices to monitor their bidirectional forwarding paths to serve for superstratum protocols. However, there is no discovery mechanism for BFD, it is notified by superstratum protocol to establish sessions. After a session is established, if no BFD control packet is received from the peer within detection time, it notifies the failure to superstratum protocol which will take appropriate measures.

4.8.2 BFD Configuration Task List

1. Configure BFD basic function
2. Configure BFD for RIP (ng)
3. Configure BFD for static route (IPv6)
4. Configure BFD for VRRP (v3)

1. Configure BFD basic function

Command	Explanation
Global Mode	
bfd mode{active | passive} **no bfd mode**	Configure the mode before establishing BFD session, the default is active mode. No command restores active mode.
bfd authentication key <*1-255*> text <*WORD*> **no bfd authentication key <*1-255*>**	Configure key and authentication character string encrypted with text for BFD, no command deletes the configured key.
bfd authentication key <*1-255*> md5 <*WORD*> **no bfd authentication key**	Configure key and authentication character string

	encrypted with md5 for BFD, no command deletes the configured key.
Interface Mode	
bfd interval <*value1*> min_rx <*value2*> multiplier <*value3*> **no bfd interval**	Configure the minimum transmission interval and the multiplier of session detection for BFD control packets, no command restores the default detection multiplier.
bfd min-echo-receive-interval <*value*> **no bfd min-echo-receive-interval**	Configure the minimum receiving interval for BFD control packets, no command restores its default value.
bfd echo **no bfd echo**	Enable bfd echo, no command disables the function.
bfd echo-source-ip <*ipv4-address*> **no bfd echo-source-ip**	Detect link fault by configuring source address of echo packets, no command deletes the configured source address of echo packets.
bfd echo-source-ipv6 <*ipv6-address*> **no bfd echo-source-ipv6**	Detect link fault by configuring source address of echo packets, no command deletes the configured source address of echo packets.
bfd authentication key <*1-255*> **no bfd authentication key**	Enable BFD authentication and configure key for interface, no command disables BFD authentication.

2. Configure BFD for RIP (ng)

Command	Explanation
Interface Mode	
rip bfd enable **no rip bfd enable**	Configure BFD for RIP protocol on the specific interface, no command disables BFD for RIP protocol.
ipv6 rip bfd enable **no ipv6 rip bfd enable**	Configure BFD for RIPng protocol on the specific interface, no command cancels the configuration.

3. Configure BFD for static route (IPv6)

Command	Explanation
Global Mode	
ip route {vrf <name> <ipv4-address> | <ipv4-address>} mask <nexthop> bfd **no ip route {vrf <name> <ipv4-address> | <ipv4-address>} mask <nexthop> bfd**	Configure BFD for the static route, no command cancels the configuration.
ipv6 route {vrf <name> <ipv6-address> | <ipv6-address>} prefix <nexthop> bfd **no ipv6 route {vrf <name> <ipv6-address> | <ipv6-address>} prefix <nexthop> bfd**	Configure BFD for the static IPv6 route, no command cancels the configuration.

4. Configure BFD for VRRP (v3)

Command	Explanation
VRRP(v3) Group Configuration Mode	
bfd enable **no bfd enable**	Enable BFD for VRRP(v3) protocol and enable BFD detection on this group, no command disables the function.

4.8.3 Examples of BFD

4.8.3.1 Example for Linkage of BFD and Static Route

Example:

Configure a static route to 14.1.1.0/24 on Switch A and configure a static route to 15.1.1.0/24 on Switch B. Both switches enable BFD detection. When the link between Switch A and Switch B is failing, BFD can detect it immediately.

Figure 4-17

Configuration procedure:

Switch A:

Switch#config

Switch(config)#interface vlan 12

Switch(config-if-vlan12)#ip address 12.1.1.1 255.255.255.0

Switch(config)#interface vlan 15

Switch(config-if-vlan15)#ip address 15.1.1.1 255.255.255.0

Switch(config)#ip route 14.1.1.0 255.255.255.0 12.1.1.2 bfd

Switch B:

Switch#config

Switch(config)#interface vlan 12
Switch(config-if-vlan12)#ip address 12.1.1.2 255.255.255.0
Switch(config)#interface vlan 14
Switch(config-if-vlan15)#ip address 14.1.1.1 255.255.255.0
Switch(config)#ip route 15.1.1.0 255.255.255.0 12.1.1.1 bfd
When the link between Switch B and layer 2 switch is failing, Switch A can detect the change of Switch B immediately, here the static routing is at inactive state.

4.8.3.2 Example for Linkage of BFD and RIP Route

Example:

Switch A and Switch B are connected and run RIP protocol, both of them enable BFD function. When the link between Switch A and Switch B is failing, BFD can detect it immediately.

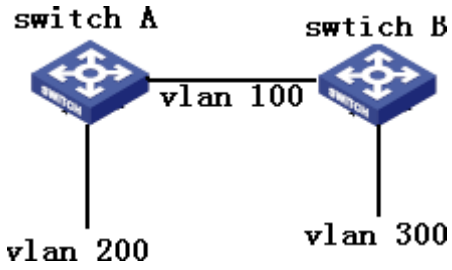

Figure 4-18

Configuration procedure:
Switch A:
Switch#config
Switch(config)#bfd mode active
Switch(config)#interface vlan 100
Switch(config-if-vlan100)#ip address 10.1.1.1 255.255.255.0
Switch(config)#interface vlan 200
Switch(config-if-vlan200)#ip address 20.1.1.1 255.255.255.0
Switch(config)#router rip
Switch (config-router)#network vlan 100
Switch (config-router)#network vlan 200
Switch(config)#interface vlan 100
Switch(config-if-vlan100) #rip bfd enable
Switch B:
Switch#config
Switch(config)#bfd mode passive
Switch(config)#interface vlan 100
Switch(config-if-vlan100)#ip address 10.1.1.2 255.255.255.0
Switch(config)#interface vlan 300
Switch(config-if-vlan300)#ip address 30.1.1.1 255.255.255.0
Switch(config)#router rip
Switch (config-router)#network vlan 100

Switch (config-router)#network vlan 300
Switch(config)#interface vlan 100
Switch(config-if-vlan100) #rip bfd enable
When the link between Switch A and Switch B is failing, BFD can detect it immediately and notifies RIP to delete the learnt route.

4.8.3.3 Example for Linkage of BFD and VRRP

Example:

When the master is failing, the backup cannot become the master until the configured timeout timer expires. The timeout is generally three to four seconds and therefore the switchover is slow. To solve this problem, VRRP uses BFD to probe the state of the master. Once the master fails, the backup can become the new master within 100 ms.

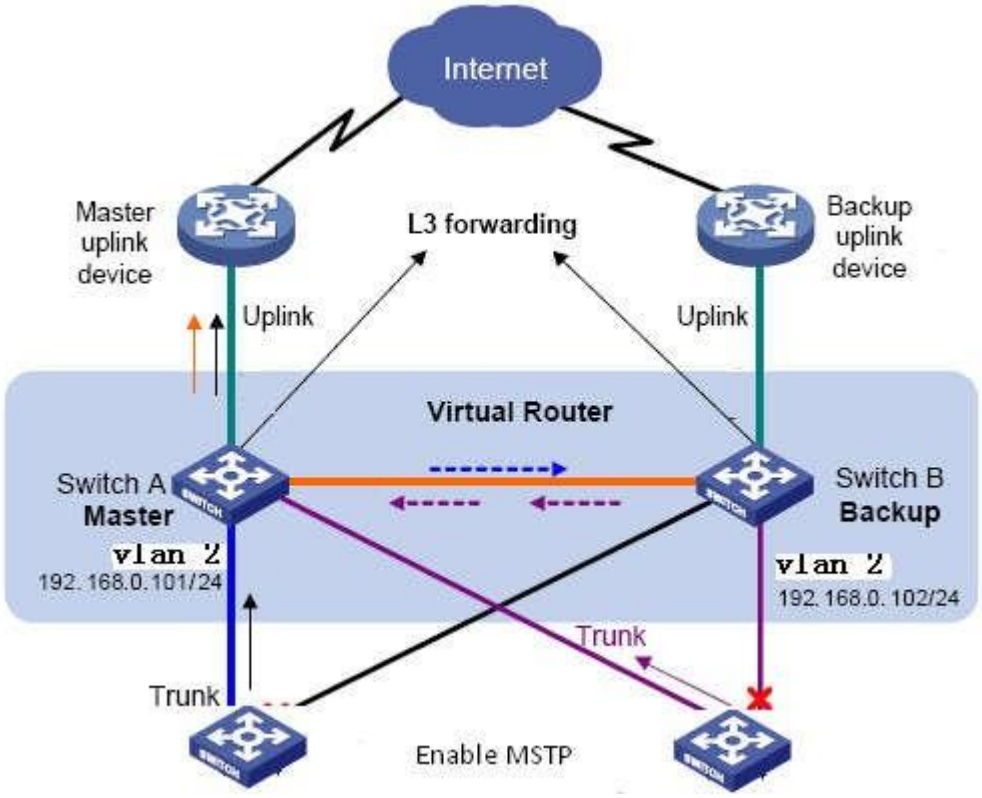

Figure 4-19

Configuration procedure:
Configure Switch A
Switch#config
Switch(config)#bfd mode active
Switch(config)#interface vlan 2
Switch(config-ip-vlan2)#ip address 192.16.0.101 255.255.255.0
Switch(config)#router vrrp 1
Switch(config-router)#virtual-ip 192.168.0.10
Switch(config-router)#interface vlan 1

```
Switch(config-router)#enable
Switch(config-router)#bfd enable

# Configure Switch B
Switch#config
Switch(config)#bfd mode passive
Switch(config)#interface vlan 2
Switch(config-ip-vlan2)#ip address 192.16.0.102 255.255.255.0
Switch(config)#router vrrp 1
Switch(config-router)#virtual-ip 192.168.0.10
Switch(config-router)#interface vlan 1
Switch(config-router)#enable
Switch(config-router)#bfd enable
```

4.8.4 BFD Troubleshooting

When the problem of BFD function happens, please check whether the problem is resulted by the following reasons:

- ☞ Check whether the route protocol neighbor is established successfully. If no route protocol neighbor is established successfully, here BFD can not process the detection.
- ☞ Check whether the configured source-ip is correct for linkage with static route, if the connectivity of IP between two peers fails, BFD can not process the detection.
- ☞ Check whether VRRP group is established successfully for linkage with VRRP protocol. If no VRRP group is established successfully, here BFD can not process the detection.

4.9 BGP GR

4.9.1 Introduction to GR

Along with network development, it requires the higher availability, so HA (High Availability) is set, namely, how to ensure packets to be forwarded and does not affect traffic operation when router control layer can not work normally.

Usually, when a router does not work normally, neighbor in route protocol layer will detect their relationship to be down, and is up soon. The process is called neighborhood shock. This shock will result the router shock that will eventually result router black hole or data passed by restarted router. Finally network availability will decrease quickly.

In order to achieve high availability, it needs upper layer route protocol to support GR (Graceful Restart). Use GR can ensure that packets can be processed or forwarded correctly when the control layer is failing.

GR can reduce route shock, resource expend consumption of control layer and improve network stability. What describe in this document is GR, which can restart BGP protocol without affecting forwarding process, and forward packets in the correct path.

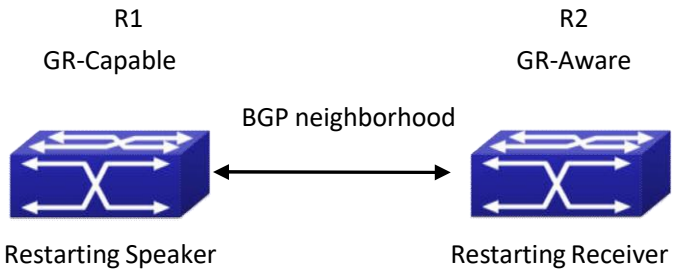

Figure 4-20 Aplication environment for GR

GR needs cooperation of GR-Capable router and GR-Aware router to complete. A restarted router is call Restarting Speaker（or GR-Restarter）, and its neighbor can be called Receiving Speaker（or GR-Helper）. Restarting Speaker is GR-Capable router while Receiving Speaker is GR-Aware router. In this way, they can complete GR. Suppose that router R1 and R2 establish BGP neighborhood, as shown in Fig 13-1, GR process can be described as:

Restarting Speaker（GR-Restarter）:

1. R1 and R2 negotiate GR capability through OPEN when establish original BGP neighbor.

2. When R1 is restarted, the route is kept in the interface board and guide forwarding continuously.

3. R1 establishes TCP connection with R2 again, it sets Restart state to 1 in BGP OPEN messages to show that this router has been restarted. At the same time, it will inform the value of restart time (it is less than Holdtime in OPEN messages) to neighbor. Additionally, it should inform neighbor what type of GR is supported.

4. After R1 is established connection with R2 correctly, it is able to receive and deal with the update information and enable selection deferral timer.

5. R1 delays the count process of the local BGP route until it receives all End-of-RIB from BGP neighbors in GR-Aware or until the local selection deferral timer is overtime.

6. Count route and send the update route. After that, it will send End-of-RIB to neighbors.

Restarting Speaker（GR-Helper）:

1. R1 and R2 negotiate GR capability with the restarted router when they establish the original neighborhood with BGP, R1 is a router that support GR-Capable.

2. When R1 is restarted, R2 may senses that TCP between R1 and R2 is cut off or cannot detect the previous state before they establish TCP connection again. If it does not detect it, go to step 4, otherwise go to step 3.

3. Keep the route sent by R1 and mark a stale label. After that, enable Restart Timer.

4. Cut off old TCP connection and deal with new TCP connection continuously. Keep the route sent by R1 and mark a stale label. After that, enable Restart Timer.

5. Establish a new neighborhood with the restarted router, delete Restart Timer and enable Stale Path Timer.

6. Before establish the new neighborhood, If Restart Timer is overtime, Restart flag does not equal 1, or there is no relevant supporting information in AFI/SAFI address family, please clear the kept route.
7. Send the route update information to the restarted router, after that, it will send End-Of-RIB label.
8. If Stale Path Timer is overtime, clear the kept route.

4.9.2 GR Configuration Task List

1. Configure whether GR capablility is supported
2. Configure whether the specific neighbor supports GR capablility
3. Configure restart-time
4. Configure restart-time for neighbor
5. Configure stale-path-time for BGP GR
6. Configure selection-deferral-time for BGP GR

1. Configure whether GR capablility is supported

Command	Description
BGP route configuration mode	
bgp graceful-restart **no bgp graceful-restart**	Enable BGP to support GR.

2. Configure whether the specific neighbor supports GR capablility

Command	Description
BGP protocol unicast address family mode and VRF address family mode	
neighbor (A.B.C.D | X:X::X:X | WORD) capability graceful-restart **no neighbor (A.B.C.D | X:X::X:X | WORD) capability graceful-restart**	Set a label for neighbor, it takes GR parameter when send OPEN messages.

3. Configure restart-time

Command	Description
BGP route configuration mode	
bgp graceful-restart restart-time <1-3600> **no bgp graceful-restart restart-time <1-3600>**	Configure BGP GR's restart-time (Receiving Speaker enables a timeout timer for a neighbor, it uses the restart-time as the timeout). A restart-time specifies the longest waiting time from Receiving Speaker finds restarting to the received OPEN messages. If Receiving Speaker does not receive OPEN messages after exceed the time, it can delete SATLE route saved by neighbor.

4. Configure restart-time for neighbor

Command	Discription
BGP protocol unicast address family mode and VRF address family mode	
neighbor (A.B.C.D | X:X::X:X | WORD) restart-time <1-3600> **no neighbor (A.B.C.D | X:X::X:X | WORD) restart-time <1-3600>**	Configure restart-time for neighbors, no command restores the default time.

5. Configure stale-path-time for BGP GR

Command	Discription
BGP route configuration mode	
bgp graceful-restart stale-path-time <1-3600> **no bgp graceful-restart stale-path-time <1-3600>**	Stalepath-time uses the default value of 360s, which is much longer than restart-time and selection-deferral-time. Because during the time from Receiving Speaker receives OPEN messages to receives EOR, it sends the initial route update and waits that the initial route update is received completely.

6. Configure selection-deferral-time for BGP GR

Command	Description
BGP route configuration mode	
bgp selection-deferral-time <1-3600> **no bgp selection-deferral-time <1-3600>**	Specify the longest waiting time that start to count selection route from the received OPEN messages to the received EOR for Restarting Speaker. If Restarting Speaker does not receive EOR after exceed the time, it can count selection route.

4.9.3 Typical Example of GR

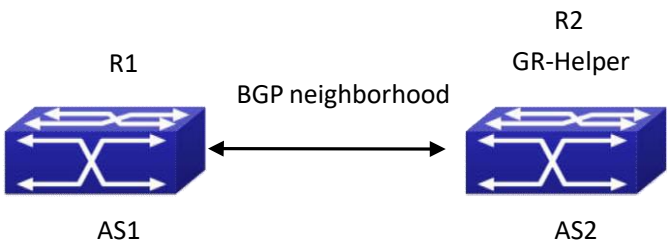

Figure 4-21 Application environment for GR

As shown in Figure 4-21, R1 and R2 establish BGP neighborhood. When they cut off the connection, BGP protocol in R2 will be in helper mode, keep route entry from R1 to R2 and restart restart-time timer. During that time, if it receives open messages from R1 or timer is

overtime, the route marked with stale in R2 will be deleted. When they establish neighborhood again, R1 will enable selection timer and wait R2 to send EOR messages or until timer is overtime, R1 is a selection route. However, after R2 receives open messages from R1, it enables STALE route timer and it will delete timer and stale route when accept EOR from R1 or timer is overtime.

R1 configures int vlan 12，ip address 12.1.1.1
R2 configuresint vlan 12，ip address 12.1.1.2
R1 configuration:
R1#config
R1(config)#vlan 12
R1(config-vlan12)#int vlan 12
R1(config-if-vlan12)#ip address 12.1.1.1 255.255.255.0
R1(config-if-vlan12)#exit
R1(config)#router bgp 1
R1(config-router)#neighbor 12.1.1.2 remote-as 2
R1(config-router)#neighbor 12.1.1.2 capability graceful-restart
R1(config-router)#bgp selection-deferral-time 120
R1(config-router)#bgp graceful-restart restart-time 60
R1(config-router)#bgp graceful-restart stale-path-time 180
R1(config-router)#exit

R2 configuration:
R2#config
R2(config)#vlan 12
R2(config-vlan12)#int vlan 12
R2(config-if-vlan12)#ip address 12.1.1.2 255.255.255.0
R2(config-if-vlan12)#exit
R2(config)#router bgp 2
R2(config-router)#neighbor 12.1.1.1 remote-as 1
R2(config-router)#neighbor 12.1.1.1 capability graceful-restart
R2(config-router)#bgp selection-deferral-time 120
R2(config-router)#bgp graceful-restart restart-time 60
R2(config-router)#bgp graceful-restart stale-path-time 180
R2(config-router)#exit

4.10 OSPF GR

4.10.1 Introduction to OSPF GR

OSPF Graceful-Restart（short for OSPF GR），is used to maintain data forwarding correctly and flow of crucial service is not interrupted when routing protocol restarts or switchover of layer 3

switches between active master and standby master. It is one of high availability technologies.

So far, the high layer 3 switches usually adopt a design for separating control and forwarding. The control module for counting routing protocol at master control board, but data forwarding module is at liner card. As a result, it will not affect data forwarding on line card when the master control board is restarted. So the device supporting GR is generally a chassis device and has two master control boards.

Since standard OSPF protocol (RFC2328) does not support GR, it will lead to flow cut off and routing surge when routing protocol is restarted or switchover between active master and standby master for various reasons. For example, as shown in below figure, when S1 occurs switchover, the neighborhood relation between S1 and S2 will lose, at that time S2 will send Router-LSA to S3 and S4 and this LSA does not include the link between S1 and S2. After S3 and S4 received LSA, they will count routing protocol again. The result will not include the link between S1 and S2. After S1 finishes the switchover, it will establish neighborhood relation with S2 and synchronize database, this action leads S2, S3 and S4 to count routing again. However, switchover of S1 will result routing shiver, which is not accepted by some networks with high requirement for performance.

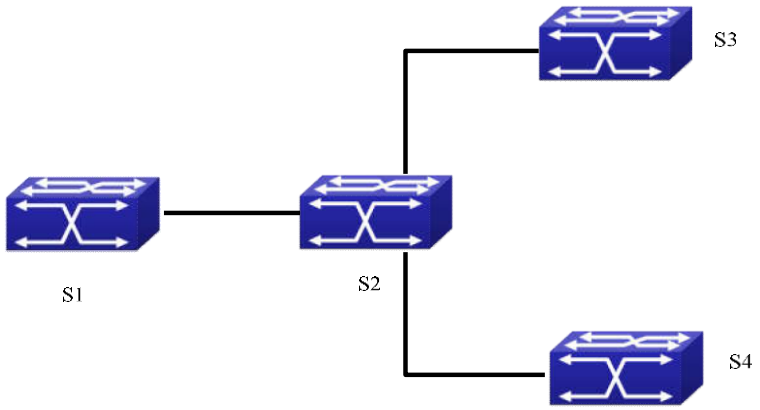

Figure 4-22 typical application scene

OSPF GR described in RF C3623 is come up for the above state. Its basic idea is that if the network topology keeps stabilization during the switchover and layer 3 switch can maintain the same forwarding list, then its neighbor can maintain their relationship, which can make the switch on its forwarding path still. If S1 and S2 support and enable GR, the liner card of S1 will keep the traffic forwarding and S2 can maintain the relationship with S1, at the same time, network topology between S3 and S4 will not be changed, furthermore, it does not need to count routing again. All of these ensure the traffic forwarding and avoid routing shiver.

Layer 3 switch can be divided into GR restarter and GR helper according to its function in GR process. GR restarter is layer 3 switch to occur the switchover between active master and standby master or restart protocol while GR helper is layer 3 switch to help GR restarter. In the above example, S1 is GR restarter and S2 is GR helper

The advantages of OSPF GR in the following:

- ☞ Increase network reliability
- ☞ Reduce the effect of routing shiver to network
- ☞ Reduce the effect to traffic and avoid that lose packets during switchover

4.10.2 OSPF GR Configuration

OSPF GR configuration task list:

1. Enable GR for OSPF
2. Configure grace-period for OSPF GR restarter (optional)
3. Configure policy for OSPF GR helper (optional)

1. Enable GR for OSPF

Command	Description
OSPF protocol configuration mode	
capability restart graceful **no capability restart**	Enable GR of specific OSPF.

2. Configure grace-period for OSPF GR restarter (optional)

Command	Description
Global configuration mode	
ospf graceful-restart grace-period <*integer*> **no ospf restart grace-period**	Configure grace period for GR restarte (The switch is used to the switchover or restart the protocol). The no command restores its default value.

3. Configure the policy for OSPF GR helper (optional)

Command	Description
Global configuration mode	
ospf graceful-restart helper max-grace-period <*integer*> **no ospf graceful-restart helper**	One of GR helper policy. Configure maximum grace period supported by helper. The no command deletes all configured helper policy.
ospf graceful-restart helper never **no ospf graceful-restart helper**	One of GR helper policy. Configure the switch can not become OSPF GR helper. The no command deletes all configured helper policy.

4.10.3 OSPF GR Example

Example:

There are for switches from S1 to S4 (They are two master control board and supports OSPF GR), they enable OSPF to implement the following functions:

1. S1 keeps traffic forwarding during the switchover, S2-S4 ensure that no routing shiver and the continuous network traffic.

2. S1 needs to finish the switchover and restart protocol within 120s, otherwise S2 will quit GR and count routing again.
3. S1 does not work as a OSPF GR Helper (S1 will not help S2 to process GR, but it will count routing again when S2 processes the switchover or restart OSPF protocol).

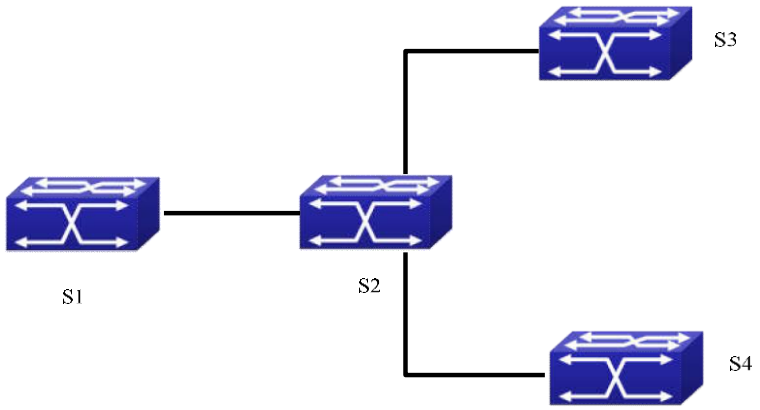

Figure 4-23 Typical application

Steps: Since the switch enables OSPF GR by default, we only need to configure the parameters and helper policy for OSPF GR. (the following configuration is relative with OSPF GR only and that of topology is omitted).

S1

S1(config)#ospf graceful-restart grace-period 120

S1(config)# ospf graceful-restart helper never

S2

S2(config)# ospf graceful-restart helper max-grace-period 120

4.10.4 OSPF GR Troubleshooting

When you have trouble in using OSPF GR, please check the following reasons:

- ☞ Whether GR restarter switch supports OSPF GR and has two main control boards, please ensure that specific GR is not disabled.
- ☞ Whether network topology is changed during OSPF GR process. When it is changed, switch may quit GR and restart OSPF.
- ☞ Please ensure all neighbors of GR restarter support GR.
- ☞ Do not modify the relevant configuration of OSPF during GR.